Universitat Politècnica de Catalunya

Departament de Llenguatges i Sistemes Informàtics

On Collective Computation

Jordi Delgado Pin

Barcelona, Setembre de 1997

Universitat Politècnica de Catalunya

Departament de Llenguatges i Sistemes Informàtics

On Collective Computation

Jordi Delgado Pin

Barcelona, Setembre de 1997

Universitat Politècnica de Catalunya

Departament de Llenguatges i Sistemes Informàtics

On Collective Computation

Certifico que la present tesi doctoral ha estat realitzada sota la meva direcció.

Barcelona, Setembre 1997.

Signat: Dr. Ricard Vicente Solé
Professor Titular del
Dept. de Física i Enginyeria Nuclear de la
Universitat Politècnica de Catalunya.

A Montse

Agraïments (provisionals)

Aquesta tesi és fruit de la col.laboració amb en Ricard V. Solé, director d'aquest treball, a qui li dec moltes de les idees aqui presents. A més, a ell i a la resta de l'original grup de Sistemas Complexos (Jordi, Barto, Sus) els dec cinc anys de fer ciència de manera, diguem, heterodoxa (i divertida!). Als nou vinguts (Claudia i Javier) els dec el tenir ben present la fascinació (i les dificultats) dels començaments. Us envejo.

Hi ha gent sense els que no seria el que soc ni faria el que faig: El meu cosí Joan Antoni i el meu oncle Manel van estar els iniciadors del meu interès per conèixer, encara que després han vingut els membres del G.E.A. i antics Sampsorites (ells saben a qui em referixo), a la colla del Pepele, el Manel, el Carmelo i Ana, el Jordi i la Saysamon (sorry, no sabré mai escriure el teu nom), el Ferran i la Sonia; i al Jordi i l'Elena, sempre aportant punts de vista originals sobre tot el que els hi comentes. Hauria de mencionar de manera especial l'amistat que m'uneix al Ferran. Vam coincidir davant d'una nevera parlant de cinema negre i des d'aleshores un munt de coses ens han passat. I n'hi ha encara un munt per venir!

A LSI hi ha una pila de gent de qui he apres moltíssim: Ton Sales, Quim Gabarró, Jacobo Toran, Ricard Gavaldà, Xavier Messeguer, Conrado Martinez, José Luis Balcazar, Felipe Cucker i segur que em deixo algú (que em perdoni, sisplau). A tots ells, de fet, els dec o bé coneixements o bé motivació (o ambdues coses). Ha estat un plaer aprendre amb ells l'autèntica informàtica, vull dir la interessant. També a LSI, no me n'hauria sortit sense l'ajut del personal d'administració, la Mercé, la Nuria, la Rita, la Judit i l'Ana.

A l'staff de l'Institut de Santa Fe li agraeixo l'hospitalitat que va demostrar els mesos de Juliol de 1996 i Juliol de 1997. M'ho vaig passar realment bé sopant amb en Stuart Kauffman i/o en Brian Goodwin i presenciant les bromes que es feien l'un a l'altre. Guy Theraulaz i Eric Bonabeau van fer que la meva estada del darrer any fos més divertida, a més de proporcionar converses força interessants, i a James P. Crutchfield li he d'agrair el temps que em va dedicar: La seva feina i les seves idees són una abundant font d'inspiració.

Aquests anys no els podria haver passat sense l'ajut afectiu dels meus pares Leandre i Marta, els meus germans Ferran i Meritxell i els meus avis Leandro, Maria, Marcelino

i Pepita. No puc deixar d'incloure la part "normalitzada" de la meva familia, el meu oncle Manel i la tia Carlota, més l'Arianna, la Cora i la Selma. I la padrineta Maria Rosa (el meu primer ordinador el va comprar ella!). Us estimo.

Finalment, a qui va dedicada la tesi, la Montse fa dotze anys que és el meu amor. Què més puc dir?

Summary

1 Introduction 1
- 1.1 Reactive DAI . . . 2
- 1.2 Self-organization and collective behavior . . . 7
- 1.3 Emergent functionality . . . 11
- 1.4 Social insects and RDAI . . . 12
- 1.5 On collective computation . . . 16

2 Noise-induced phenomena and Fluid Neural Networks 19
- 2.1 Self-synchronized behavior in ant colonies . . . 19
- 2.2 Fluid Neural Networks (FNN) . . . 24
  - 2.2.1 Definition . . . 24
  - 2.2.2 Analysis of individual behavior . . . 27
  - 2.2.3 Collective behavior in FNNs: Oscillations . . . 29
- 2.3 Phase transitions in FNNs . . . 33
  - 2.3.1 Noise-induced transitions: A brief introduction . . . 33
  - 2.3.2 Order parameter for FNNs . . . 36
- 2.4 Mean field approximation of ρ_c . . . 39
- 2.5 Self-synchronization and task fulfilment . . . 47
- 2.6 Conclusion . . . 52

3 Computation in collectives of simple agents 55
- 3.1 Dynamics and computation . . . 55
  - 3.1.1 Dynamics, but... . . . 56
  - 3.1.2 ...where is the computation? . . . 64
- 3.2 Universal computation in FNNs . . . 67
  - 3.2.1 Matrix Λ_0, chemical switch . . . 71
  - 3.2.2 Matrix Λ_F, nonzero threshold . . . 72
  - 3.2.3 Matrix Λ_F, two chemicals . . . 73
  - 3.2.4 Parameter space . . . 74
- 3.3 Conclusion . . . 75

4 Collective computation: Theoretical approach 79
- 4.1 Computational mechanics . . . 79
- 4.2 RDAI and dynamical systems . . . 87

4.3 Globally coupled maps 88
4.3.1 Universality in Discrete Dynamical Systems 88
4.3.2 Coupled maps 91
4.4 Individual vs. collective 96
4.4.1 Problem statement 96
4.4.2 Collective-induced computation 97
4.5 Conclusion 102

5 Overview and prospects **105**
5.1 Summary 105
5.2 Prospects 108

6 References **115**

Chapter 1

Introduction

What is it that governs here,
that issues orders, foresees the future...?
M. Maeterlinck (1927)

It has been long since the french ethnologist Claude Lévi-Strauss pointed out the unavoidable liaison that any psychological study has to have with social facts (Lévi-Strauss, 1962). He was not the first one to make such considerations, and undoubtedly he was not the last: There are lots of works on the social basis of intelligence. In spite of that, Artificial Intelligence (AI), from its very beginning, has been trying to get intelligent machines out of solitary machines (McCorduck, 1979). Whether they were expert systems, neural networks or more general "cognitive architectures" (Russell & Norvig, 1995) the interaction with a "social" environment was neglected in most studies. In order to *dés-individualiser* AI it was necessary to introduce those social considerations, and problems like coordination among agents or cooperative problem solving began to be taken into account in the AI community, giving birth to the subfield of AI now known as Distributed AI (DAI) (see Bond & Gasser, 1988, for classical works in DAI and Bonabeau & Theraulaz, 1995 chap. 5 for a summary).

According to Castillo & Quintanilla (1991), DAI might be defined as *the cooperative solution of problems by a decentralized and loosely coupled collection of agents.* This definition is quite interesting since it reveals some important points we want to emphasize. DAI systems are systems composed by agents displaying sophisticated behaviour, such as reasoning or planning, where agents spend more time computing and much less time in communicating and interacting. Agents, in classical DAI systems, are viewed as deliberative; accordingly, agent theory is based on logics of knowledge and belief or, more generally, on some sort of logical framework (see the first chapter in Wooldridge & Jennings, 1995 for a review of agency theory).

Besides agent sophistication (or perhaps because of it), one of the main problems in DAI is that of cooperation and coordination, essentially, that of interaction. Several protocols have been designed to deal with interactions among agents, such as the contract net, where a task is distributed among the nodes of a net (each node may be seen as a separate agent, either a manager or a contractor) via a sophisticated mechanism called negotiation (see the original paper by Davis & Smith, 1983). It should be clear, then, why interaction must be minimised when agents perform the core of the computation and why there is a loose coupling emphasis in the above definition of DAI.

On the other hand, the initial psychological influence on AI, based mostly on the "mentalist" school where symbol manipulation and mental representations were central concepts, has been recently questioned in favour of a revival of some works of the "cybernetic age" (for example, Grey Walter's turtle, see McCorduck, 1979), where reactive creatures were built with neither representations nor symbols, only stimulus-response activity. Brooks' subsumption architectures (Brooks, 1991) are representatives of current research in what is nowadays known as "Behavior-Oriented" AI (BOAI). BOAI was a quite natural reaction to old critiques to AI (Dreyfus, 1992), critiques that pointed out the disembodied nature of the systems built by AI in the sixties. Again, a lack of interaction of the AI systems, this time with an environment, was the cause of a reconceptualization of AI research.

Taken separately, DAI and BOAI are only two of a long list of subfields of AI (knowledge representation, machine learning, neural networks, genetic algorithms and a long etcetera) but the study of collectives of reactive, simple agents gives rise to results which are extremely interesting not only to AI, but also to physics (strongly interacting many body systems), biology (models of swarm behaviour) and complex systems research (Nicolis & Prigogine, 1977, 1989). This is called *Reactive DAI.*

1.1 Reactive DAI (RDAI)

Now, let us assume that our agents are not sophisticated, but quite simple. Neither planning nor knowledge representation will be properties of these agents. The departure point in RDAI is the *reactive* agent. These agents get some input from an environment (other agents are elements of an agent environment) and *react* in a pre-programmed way to these stimuli. It is as if it had some sort of stimulus-reaction look-up table. Thus, interactions might not be much complicated, because of the sparse information processing capabilities of individual agents. Interaction among

agents may be direct (physical contact or signal synchronization) or indirect (by means of signals left on a common environment), but they must be in the form of stimuli to other agents. Hence, not much information processing is involved in the process of interaction. Furthermore, interaction, either direct or indirect, is often of a *local* nature, that is, the agent is not able to perceive stimuli beyond some well defined (though problem dependent) range from its location in the environment. There are also situations where no interactions occur, but an implicit cooperation emerges, even if each agents follows individual goals (Mataric, 1994).

It is nevertheless true that *collectives* of those agents are able to perform quite complex tasks, tasks whose difficulty is far beyond individual capabilities, as we will see below. But, if agents do not realize the computation required to solve a given task, who is doing that task? The answer is that *the collective* gets the task done: The system as described will possess some sort of configuration (the state of all agents plus other relevant properties, for example their location in the environment, if space is important) that will change with time, as the system evolves. It is this *global* or *system level* configuration that will solve the problem, since, in some problem-dependent sense, the global configuration of the system *is* the solution.

Let us make concrete this general description with a detailed example of a system of elementary agents capable of choosing between two strategies (see Ceccatto & Huberman, 1989, though the description given here is slightly different). Assume N agents $e_i \in \{-1, +1\}$ (if $e_i = -1$ the agent e_i is using the strategy 1). What is of interest to us is some measure of the number of agents using strategy 1, from which we can compute trivailly the number of agents using strategy 2; this measure will be $\mu(\mathbf{e}) \in [-1, +1]$

$$\mu(\mathbf{e}) = \frac{1}{N} \sum_{i=1}^{N} e_i \tag{1.1}$$

Two functions $G_1(\mu)$ and $G_2(\mu)$ will give the payoff of using the i-th strategy ($i = 1, 2$). There are diverse possible definitions of G_i, depending on the problem we want to deal with, though we will keep them undefined since we want a model as general as possible. The mechanism of strategy selection will be a probabilistic one, where $P(e_i = -1)$ will be the probability that agent e_i perceives strategy 1 to be better than strategy 2. For the sake of analytical tractability it will be defined as

$$P(e_i = \pm 1) = \frac{1}{2}(1 \pm \tanh(\beta[G_1(\mu) - G_2(\mu)])) \tag{1.2}$$

The description of the system is already complete. Let us emphasize the simplicity of the agents (± 1 variables) and of the mechanism of decision, mechanism that

depends on the state of all other agents through the $\mu(\mathbf{e})$ function, i.e. interactions are *not* local. The system, as defined, may be used to model diverse situations, such as resource allocation or the competition between cooperation and crowding (see Huberman, 1988). Now we must study the *dynamical behaviour* of the quantity of interest, that is, the mean of $\mu(\mathbf{e})$ (denoted by $< \mu(\mathbf{e}) >$), since there are two sources of randomness: The probabilistic behavior of decision and the initial condition. We can define a probability transition $W(\mathbf{I} \to \mathbf{J})$ between two different global states $\mathbf{I} = (e_1^I, \ldots, e_N^I)$ and $\mathbf{J} = (e_1^J, \ldots, e_N^J)$

$$W(\mathbf{I} \to \mathbf{J}) = \begin{cases} 0 & \text{if } |\mu(\mathbf{I}) - \mu(\mathbf{J})| > \frac{2}{N} \\ P(e_{i^*}) & \text{if } |\mu(\mathbf{I}) - \mu(\mathbf{J})| = \frac{2}{N} \end{cases} \tag{1.3}$$

where i^* is the agent with different state in $\mathbf{I}$ and $\mathbf{J}$. Finally, with $W(\mathbf{I} \to \mathbf{J})$ we can find the equation of the temporal evolution of the probability distribution $P_t(\mathbf{e})$

$$P_{t+\Gamma}(\mathbf{I}) = \sum_{\{\mathbf{J}\}} W(\mathbf{J} \to \mathbf{I}) P_t(\mathbf{J}) \tag{1.4}$$

also known as master equation (Gardiner, 1983). The mean $< \mu(\mathbf{e}) > (t)$ is taken with respect to $P_t(\mathbf{e})$. Some standard approximations (not detailed here, see Amit, 1989, pp. 148-150) can be performed in order to get the desired result

$$\Gamma \frac{d}{dt} < \mu(\mathbf{e}) > (t) = - < \mu(\mathbf{e}) > + < \tanh(\beta[G_1(\mu(\mathbf{e})) - G_2(\mu(\mathbf{e}))]) > \tag{1.5}$$

The term $< \tanh(\ldots) >$ is difficult to compute so a standard procedure is applied, the "mean field" approximation, known to be exact in the thermodynamic limit $N \to \infty$ (Binney *et al.*, 1992)

$$\Gamma \frac{d}{dt} < \mu(\mathbf{e}) > (t) = - < \mu(\mathbf{e}) > + \tanh(\beta[G_1(< \mu(\mathbf{e}) >) - G_2(< \mu(\mathbf{e}) >)]) \tag{1.6}$$

Several conclusions might be obtained from this equation. The dynamics is of the form (if $\mu \equiv < \mu(\mathbf{e}) >$)

$$\frac{d\mu}{dt} = -\frac{\partial \Omega(\beta, \mu)}{\partial \mu} \tag{1.7}$$

therefore it is governed by the zeroes of $g(\beta, \mu) = \partial \Omega(\beta, \mu)/\partial \mu$, so, given an initial condition, the system will relax to the closest zero of g, though at longer times the system will reach the global minimum of Ω (see the discussion on the slaving

principle below and Binney *et al.*, 1992). The point is that local minima of Ω (also called "metastable" states) will be the "real" stable states if the system is large, therefore non-optima states may "trap" the system (see Ceccatto & Huberman, 1989, and Huberman, 1988, for a complete analysis of the problem).

We are not so interested in the details of the example as we are interested in how the example illustrates the general discussion on RDAI. We easily see that the dynamics of the quantity $< \mu(\mathbf{e}) >$, dependent on the *global* state $\mathbf{e} = (e_1, \ldots, e_N)$ of the system, is the relevant statistic parameter of interest, though is not trivially infered from the individual dynamics. Furthermore, the optimal performance of the system depends on a complex dynamics, being the solution strongly related with the notion of "attractor" of a dynamical system. Applying this model to a concrete problem, the optimal solution would be attained (depending on how G_1 and G_2 are defined) without neither central control nor sophisticated interaction strategies. All these features make the analysis (whenever possible) of the dynamical system underlying the agents and their interactions the core of the research in RDAI, adding new techniques from other research fields (essentially statistical physics and dynamical systems theory) to the AI toolbox. This is clearly seen, to cite a few instances, in neural networks research (Amit, 1989) and the computational ecologies paradigm of DAI (Huberman, 1988). However, there are more complicated cases where the study must rely on computer simulations, or, often, on a combination of analysis and simulation, though the principles of design are the same as those discussed above (see Beni & Hackwood, 1992; Bonabeau & Theraulaz, chapters 6 and 7, 1995; Deneubourg *et al.*, 1991 and 1992; Dorigo *et al.*, 1996; Goss & Deneubourg, 1992; Gutowitz, 1993; Kephart *et al.*, 1990; Kube & Zhang, 1993; Mataric, 1994; Sugawara & Sano, 1997). Let us slightly detail some examples.

- *Collective Algorithms*: As a representative example of the "collective of simple agents" view of problem solving we will detail a collective solution to the Travelling Salesman Problem (TSP, Colorni *et al.*, 1992; Dorigo *et al.*, 1996). Consider a set of N cities $C_1, \ldots, C_N$ and distancies between cities $d(C_i, C_j)$ for each pair (C_i, C_j). A solution to the TSP is a permutation π of cities such that the quantity

$$\sum_{i=1}^{N-1} d(C_{\pi(i)}, C_{\pi(i+1)}) + d(C_{\pi(N)}, C_{\pi(1)})$$

  is minimized. The collective is a set of M agents scattered among the cities, where $b_i(t)$ is the number of agents in the i-th city.

The algorithm is as follows: An agent in city i decides, in the time interval between t and $t+1$, to go from city i to city j with certain probability $p_{ij}(t)$

$$p_{ij}(t) = \begin{cases} \frac{[s_{ij}(t)]^{\alpha} d_{ij}^{-\beta}}{\sum_{j \in \text{allowed}} [s_{ij}(t)]^{\alpha} d_{ij}^{-\beta}} & \text{if } j \in \text{allowed} \\ 0 & \text{otherwise} \end{cases}$$

depending on the distance d_{ij} (whose inverse is called *visibility*) and the *intensity of trail* $s_{ij}(t)$ on edge (i,j) at time t. The *trail* is a kind of evanescent "formal chemical substance" that an agent modifies once it has visited all the cities (see Colorni *et al.*, 1992, for details on trail dynamics). Each agent is forced to go from a city i to those "allowed" j, that is, those cities not visited in the current tour. Once it has visited all N cities the tour starts again from a "free" state and the trail in each edge (i,j) visited is modified. α and β are parameters that allow the user to control the importance of trail versus visibility. This algorithm is cycled for all agents NC times and at the end the shortest tour found is the solution to the instance of the TSP problem.

The results of this algorithm are quite satisfactory. We only mention that, with the best parameter values, the algorithm finds solutions better than the best found with genetic algorithms and it is very fast in finding good enough (though not optimal) solutions. Furthermore, once found the best parameters (α and β) for problems with a certain dimension, the parameters are little sensitive to increasing the dimension of the problem (see Colorni *et al.*, 1992, for details on results).

The algorithm makes clear some points already mentioned above. Each agent is simply a stochastic process deciding which city to go from the city where it is. It is not able to perceive a global situation in which to choose the best next city in its tour. In this sense it only uses local information. Besides, information processing capabilities are quite limited, since the only thing it is able to do is to choose randomly a city to go and to remember which cities it has passed during the current tour. It is quite robust since the removing of a certain fraction of agents does not affect the result, though it delays the finding of the solution.

- *Collective robotics*: Another instance of collectives of simple interacting agents are groups of robots cooperating to solve some given task. Examples of tasks solved with groups of robots are finding paths to a certain location (Goss &

Deneubourg, 1993), pushing a box (Kube & Zhang, 1993) and clustering objects scattered in an arena (Deneubourg *et al.*, 1991). In this case there are some added difficulties, due to the physical embodiment of robots (though plenty of collective robotics experiments are performed by means of computer simulations). This embodiment is, in fact, an advantage since some collective behavior is often guided by physical or temporal constraints in the environment, allowing the system to act as if it had some global representation of that environment.

An example of implicit cooperation are the five robots of Kube & Zhang (1993) that, without any interaction but with an explicit non-interference command, pushed a box that cannot be pushed by a single robot, due to the weight of the box. To get the task done the robots must locate the box, move toward it without collisions, distribute along a side and push. Each robot was built according to Brooks' subsumption architecture (Brooks, 1991). Using a implicit cooperation strategy, the robots succeeded in pushing the box (see Kube & Zhang, 1993 for details). Another interesting experiment was to build robots that cluster objects. Initially 81 objects were scattered over an arena and the robots (again built according to Brooks' subsumption architecture) might take some, depending on a microswitch able to detect a certain quantity of objects pushed by the robot. These robots group the objects in little clusters that grow until they cannot be pushed by any robot. When there is no objects in little clusters to be moved by robots the task is done. This is also a case of implicit cooperation (Bonabeau & Theraulaz, 1995, pp. 198-201).

1.2 Self-organization and collective behavior

Up to now we have roughly described the systems that are the subject of RDAI research and we have also seen some quite different examples of systems that properly fall among the problems dealt with by the RDAI point of view. However, which are the reasons behind such *self-organized* behavior? It does not exist a general answer to this question, since quite different mechanisms have been found underlying different self-organized systems (Solé *et al.*, 1996). However, one of them, the *slaving principle* (Haken, 1977), is quite general and it underlies a large class of collective systems (the connectionist systems, see Millonas, 1992), so we will digress a little on it. Let us introduce the following system of two nonlinear ordinary differential equations

$$\frac{du_1}{dt} = -\gamma_1 u_1 - \alpha u_1 u_2 \tag{1.8}$$

$$\frac{du_2}{dt} = -\gamma_2 u_2 + \beta u_1^2 \tag{1.9}$$

In the absence of the system 1.8 we assume that system 1.9 is damped, $\gamma_2 > 0$. Let us analyze the system in the case $\gamma_2 \gg \gamma_1$, that is, we have a separation of time scales, since in absence of nonlinear terms the characteristic time is $1/\gamma_i$, so what we are assuming is that the system 1.9 gets to the equilibrium point before system 1.8 and 1.8 only "perceives" the equilibrium of 1.9. If we put $du_2/dt = 0$ (to get the fixed points of 1.9) then

$$u_2(t) \approx \gamma_2^{-1} \beta u_1^2(t) \tag{1.10}$$

and we say that 1.9 is *slaved* by 1.8. Substituting 1.10 in 1.8 we get

$$\frac{du_1}{dt} = -\frac{\partial V(u_1)}{\partial u_1} \tag{1.11}$$

where

$$V(u_1) = \frac{1}{2}\gamma_1 u_1^2 + \frac{1}{4}\frac{\alpha\beta}{\gamma_2} u_1^4 \tag{1.12}$$

Now we must solve just one nonlinear system because of the slaving. In fact, the "slow" variable u_1 governs the dynamics being followed by the "fast" variable u_2, in this sense u_1 is called the *order parameter*[1]. The parameter γ_1 determines the dynamics of 1.11 as we can see in figure 1.1. It can be shown (Haken, 1977, pp. 194-200) that if $\gamma_1 > 0$ the unique solution is $u_1(t) = 0$, and thus also $u_2(t) = 0$, but if $\gamma_1 < 0$ the steady state solution of 1.11 is

$$u_1 = \pm m_0 = \pm \left(\frac{|\gamma_1|\gamma_2}{\alpha\beta} \right)^{1/2}$$

therefore $u_2 \neq 0$. This is called *symmetry breaking* since which steady state will be the final one depends on the initial condition. Figure 1.1 also makes clear why fluctuations are of paramount importance near transitions ($\gamma_1 = 0$ in our case). Switching between attractors is easier when barriers between them are small, so fluctuations may have a greater effect around critical points. This will be the underlying idea behind the work introduced in chapter 3.

This discussion with the simple systems 1.8 and 1.9 can be generalized (Haken, 1977, 1988) to systems of greater dimension. In that case, a few variables may slave all the rest, reducing the effective dimension of the system and allowing an

[1]Not to be confused with the somewhat relaxed notion of order parameter we will use in section 2.3.2 that, though related, is not identical, see Binney *et al.*, 1992

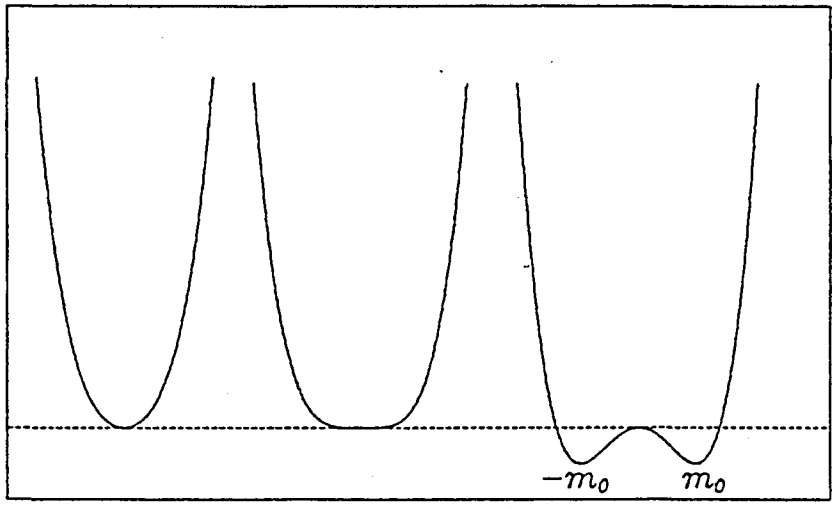


Figure 1.1: $V(u_1)$ for (left) $\gamma_1 < 0$, (middle) $\gamma_1 = 0$ and (right) $\gamma_1 > 0$. Small barriers around the critical point (middle) make switching possible with small fluctuations.

analytical treatment. When the number of order parameters is small, several tools from the theory of nonequilibrium phase transitions can be applied. In particular, the stochastic description of nonlinear dynamical systems allows us to explore the probabilistic nature of complexity. This stochastic description is not only more close to the real world, when fluctuations are unavoidable, but plays a constructive role in many instances. Specifically, noise together with nonlinearities can be a source of order. In some cases (see chapters 2,3), noise is a key ingredient for collective computation.

In this context, if a typical order parameter ξ equation is

$$\frac{d\xi}{dt} = \phi_\mu(\xi) = \lambda\xi - \beta\xi^3 + \eta(t) \tag{1.13}$$

where μ stands for the parameters λ and β, and $\eta(t)$ is a *white* noise term with

$$< \eta(t)\eta(t') >= Q\delta(t - t')$$

Assuming that $\eta(t)$ is Gaussian distributed, a Fokker-Planck (Haken, 1977) equation can be obtained:

$$\frac{\partial f(\xi,t)}{\partial t} = \frac{\partial}{\partial \xi}\left[(\lambda\xi - \beta\xi^3)f(\xi,t)\right] + \frac{Q}{2}\frac{\partial^2}{\partial \xi^2}f(\xi,t) \tag{1.14}$$

Now the stationary state $f_s(\xi,t)$ is obtained from

$$\frac{\partial f(\xi,t)}{\partial t} = 0$$

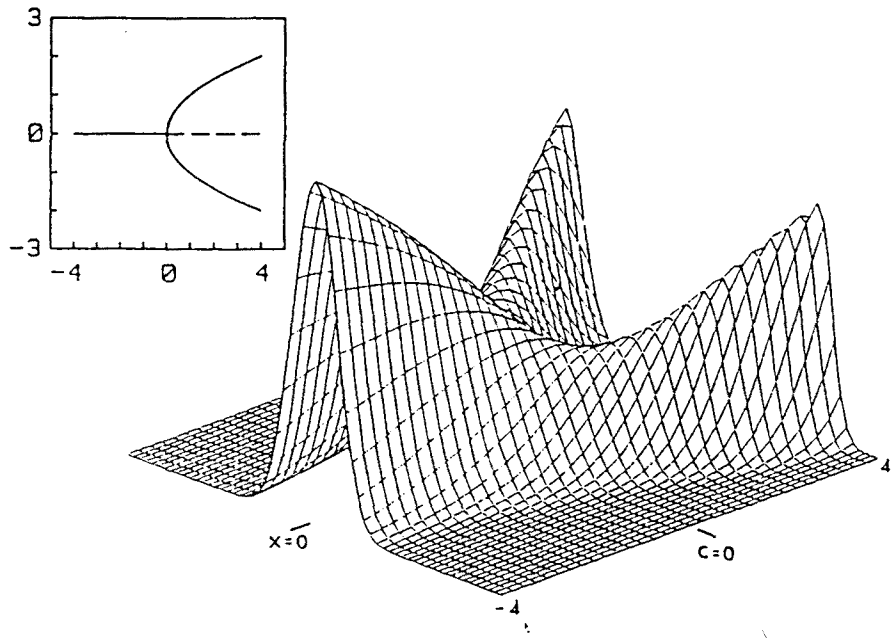


Figure 1.2: Stationary probability distribution of the order parameter 1.15. *Inset*: Bifurcation diagram for the deterministic system (see text). The stable solution $\xi^* = 0$ becomes unstable at $\lambda = 0$ leading to two new branches. The dashed line stands for the now unstable branch.

which leads to

$$f_s(\xi, t) = N \exp\left[Q^{-1}\left(\lambda\xi^2 - \frac{1}{2}\beta\xi^4\right)\right] \tag{1.15}$$

Very often, these equations are written in terms of a stochastic potential $V_\mu(\xi)$

$$f_s(\xi, t) = N \exp\left[Q^{-1}\int^{\xi} \frac{\partial V_\mu(\xi')}{\partial \xi'} d\xi'\right] \tag{1.16}$$

An example of the stationary distribution defined by 1.15 is shown in figure 1.2, when the λ parameter is varied from negative to positive values. We can see that the system moves from a single-hump distribution (centered at $\xi^* = 0$) to a distribution with two peaks, corresponding to the new alternative branches.

This approach to nonequilibrium systems has been very successful in the analysis of physical systems like lasers or chemical reactions far from equilibrium (Haken, 1977). And the concept of order parameter and phase transition has been applied to other, more complex problems, like the dynamics of the human brain (Kelso, 1995) and the construction of the so-called synergetic computers (Haken, 1988). In most of these examples the stochastic formalism involves a probabilistic description of the underlying deterministic solutions. Two different problems (both considered in

this thesis in relation with computation) are, however, not included in the previous description: Noise-induced phase transitions and adaptive systems. The first problem deals with models where fluctuations are able to generate drastic changes in the macroscopic behavior of the system, *even* outside the neighborhood of a deterministic instability point (see chapter 2). The second deals with a deeper problem, where agents interact locally through a field (the order parameter) eventually leading to emerging patterns and new attractors (chapter 3).

Why is this approach relevant to collective computing systems? Essentially because many models of computation in natural systems use the metaphor of attractors and also because noise is an unavoidable part of reality. A stable attractor (the minima of the potential $V_\mu(\xi)$) can easily be identified as the final solution for a given problem. But noise, far from being an undesirable component of this picture, can also act as a useful part in information gathering and processing. If the system is close to instability points, switching between attractors can take place (as a consequence of amplification of small inputs) and so information processing is allowed. Far beyond this point, a different possibility emerges if the creation of new attractors is available. This possibility, as explained in chapter 3, makes likely a robust information processing where noise, instead of a disturbing component, becomes an essential ingredient.

Finally, in collective systems it is frequent to find different parameters with dynamics occurring at quite different time scales, a good example would be the dynamics of synapses and the dynamics of formal neurons in neural networks (Amit, 1989). The individual neurons are the "fast" variables, reaching the attractor quite quickly. The changes in synapses (the "slow" variables) will depend on the result of the dynamics, that is, the attractor. So, synapses are almost constant values with respect to neurons and neurons are always in a steady state with respect to synapses. As we will see in section 3.1, neural networks and, more generally, connectionist systems (Millonas, 1992, 1994) are a metaphor useful to model collective phenomena, and the slaving principle will be a useful tool to solve collective systems.

1.3 Emergent functionality

Collective systems with functional properties, that is, with useful behaviour (useful in some sense, either for the system itself or for some external observer), display what is called *emergent functionality*: "Emergent functionality means that a function is not achieved directly by a component or a hierarchical system of components,

but *indirectly*[2] by the interaction of more primitive components among themselves and with the world" (Steels, 1991). Systems with emergent functionality are advantageous when compared with classical DAI systems, because of some characteristic properties, such as

- *Fault tolerance*: The removal or failure of some components of the system is dealt with easily in these systems since the extreme simplicity of the individual agents and the overall mechanisms the collective uses to solve the problem are able to compensate any moderate loss in components reliability (Steels, 1991; Forrest, 1990). This is also called *graceful degradation.* Instead, systems with few sophisticated agents cannot lose any of them without losing a considerable amount of computational power.

- *Simplicity*: The agents composing these systems, as stated above, are quite simple, so they are easily built. Deliberative agents need large amounts of information processing, since they are far from trivial knowledge systems with sophisticated interaction strategies, being much more difficult to construct. This property of simplicity is one of the reasons of a methodological difference between classical DAI and the much more practically oriented RDAI.

However, there are also some drawbacks, being the main one the far from solved problem of *design* (Mataric, 1994). As far as we know, all the existing systems with emergent functionality are either systems built from analogies with natural processes (wasps, ants, bees and termites, see section 1.2) or systems with a built-in learning algorithm in such a way that emergent behaviours are not programmed but learnt. Currently it is not known how to design, *in general*, an emergent functional behaviour from a given collective of reactive agents, neither is the relation among individual and collective capabilities (but see chapter 4).

1.4 Social Insects and RDAI

As a paradigm of swarm intelligence, and of complex systems in general, social insects have been a not yet exhausted source of multiple ideas, algorithms, mechanisms, etc... The systems RDAI practicioners aim to understand and build have been in our biosphere for more than 100 million years, so nature seems to have already solved our questions about strongly interacting simple agents. All the ants, some families of wasps and bees and the entire order of *Isoptera* (termites) are what entomologists call *eusocial.* To belong to the class of eusocial species three biological

[2]Italics are mine

traits should be present: Adults caring for the young, two or more generations of adults living together and a reproductive division of labour (Hölldobler & Wilson, 1990). From now on we will restrict our discussion to ants (needless to say that ants are representative enough of *all* kinds of social behaviour in insects). First we could ask how rare is social behaviour in nature, perhaps it has been an evolutive accident surviving only in few places, or an early error being corrected by the introduction of more individual-based species. Fortunately both statements are wrong: Social behaviour evolved from small societies of individuals, that is, ant societies with a huge number of highly cooperative individuals are the youngest, from an evolutionary point of view. With respect to their ubiquity, ant species can be found all over the world, particularly in tropical forests; let us quote Hölldobler & Wilson (1990) concerning estimates of diversity and abundance of ants:

> "All together, theses creatures seem likely to constitute half or more of the insect biomass. Consider the following disproportion: Only 13.500 species of highly social insects are known (9.500 of which are ants) out of a grand total of 750.000 insect species that have been recognized to date by biologists. Thus, more than half the living tissue of insects is made up of just 2 percent of the species, the fraction that live in well-organized colonies.(...) At the risk of oversimplification, we envisage an overall pattern of ants and termites at the ecological center, solitary insects at the periphery."

So it is quite clear that social behaviour is an advantageous strategy in environments with a certain degree of unpredictability, such as real ecosystems. It is not strange, then, the interest of artificial systems designers in social insects.

There is a large variety of interesting behaviors in ant colonies, behaviours that are attained collectively without central command. The queen, contrary to what is popularly thought, does not control role the colony, she has essentially a reproductive function. In fact, queenless colonies are not rare (see section 2.1). Examples of well-known collective functions performed by ants are task allocation, collective decision, collective sorting (all three will be detailed in section 3.1.1), foraging, patrolling and brood care, though there are more surprising behaviours in some ant species, such as propaganda, slavery, decoding, mimicry, Trojan horses (!) and highwaymen. All these striking behaviors are performed by quite simple individuals with no more than 40 "behavioral categories" interacting by means of simple mechanisms, essentially chemical communication (pheromones) and physical contact, though sound may also be used in some species (Hölldobler & Wilson, 1990).

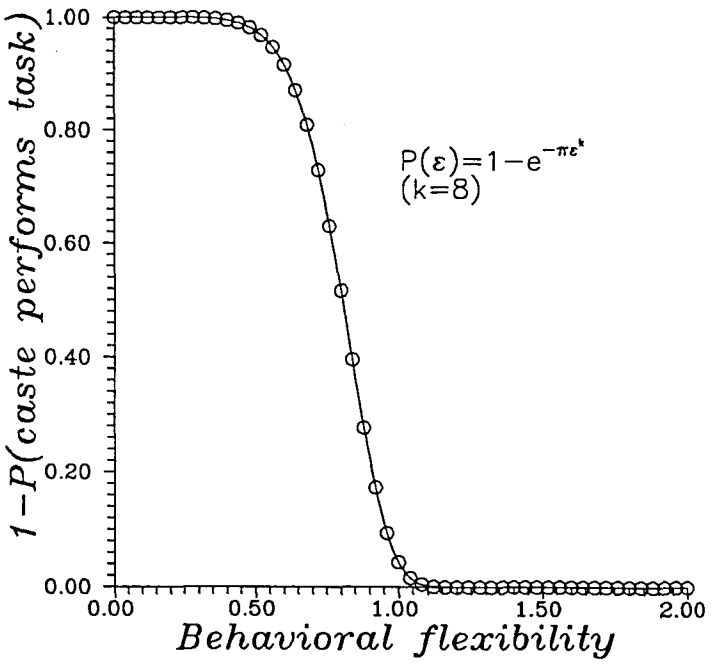


Figure 1.3: Phase transition in global performance as a function of individual flexibility ϵ. The parameter $\xi(\epsilon) = 1 - P(\epsilon)$ is a measure of how likely is that the individual remains in the same state, without switching to other tasks.

A large body of theory on optimization in ant colonies has been developed in the last decades (Oster & Wilson, 1978). These studies deal with the general problem of how to perform a given set of tasks with a given number of casts (i.e. specialized morphological types). The main object of analysis are macroscopic integrals of the kind

$$\mathcal{F}(t) = \int \eta(s) F(s, \hat{s}, t) ds d\hat{s}$$

which include the distribution of castes $\eta(s)$ together with some quantity $F(s, \hat{s}, t)$ that relates how a given benefit $\hat{s}$ is obtained for a given s. However, nonlinearities often lead to counterintuitive results, such as the relationship between behavioral flexibility (say, ϵ) and the probability of task performance $P(\epsilon)$. It can be shown that a nonlinear relation like $P(\epsilon) = 1 - e^{\pi\epsilon^k}$ is at work, k being the number of different characteristics describing a given caste. If we take $\xi(\epsilon) = 1 - P(\epsilon)$ as our macroscopic description (ξ measures how difficult is to switch from a given task to another) the plot of $\xi(\epsilon)$ strongly resembles that of a phase transition in physics. In figure 1.3 we show $\xi(\epsilon)$ (for $k = 8$) as a function of flexibility. We can see that as ϵ grows beyond a critical threshold, a sudden change is observed. This transition tells us that behavioral flexibility and the capacity to cooperate can sharply increase the range of tasks to which an individual can contribute. A consequence of this

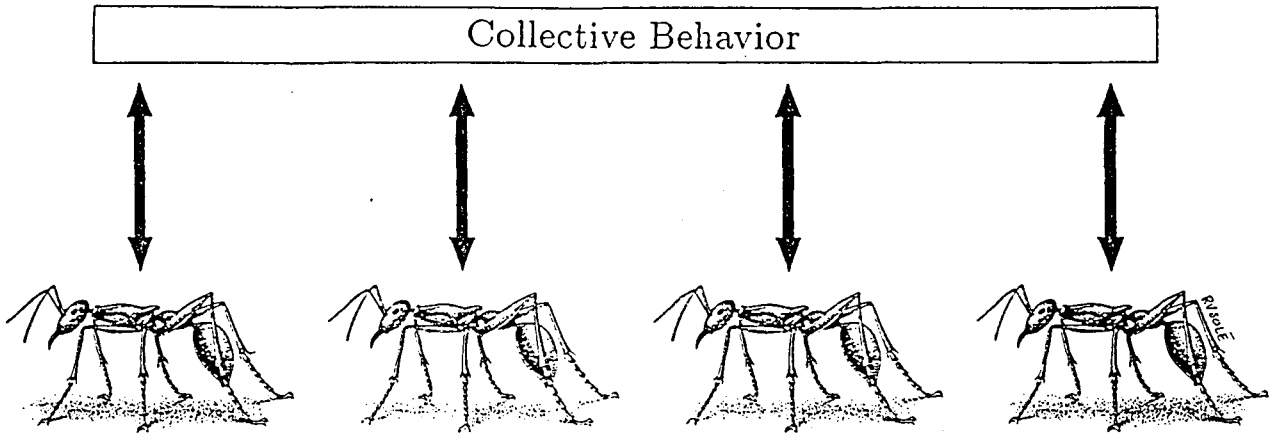


Figure 1.4: Individuals create the field that governs their behavior. This stigmergic mechanism is understood thanks to a separation of time scales, the slaving principle, that helps to explain how ants organize *collectively* their individual behavior

nonlinear relation is that a very small change in behavioral flexibility may result in a shift in the colony optimum from monomorphism to polymorphism. Quoting Oster & Wilson (1978)

> "The phenomenon is reminiscent of phase transition curves in physics, which characterize sudden condensations, shifts from order to disorder, and other abrupt transitions"

Indeed, this transition is known in physics as a second order phase transition (Binney *et al.*, 1992) and the Oster-Wilson analogy is confirmed by observations in many ant species. In very closely related species a disparate degree of polymorphism is observed, consistently with the previous scenario (Oster & Wilson, 1978).

This is just an example from real ant colonies of how important phase transitions can be. In a related context, it is not surprising that recent developments in AI involve the use of phase transitions to study NP-complete problems (Hogg & Huberman, 1987). Since phase transitions are a well-known instance of collective behavior, they will play a key role in this thesis.

Underlying all those behaviors there is what is of principal interest to RDAI, that is, the *mechanisms* of functioning, the individual rules behind the collective performance. In particular, there is a mechanism that has been the main source of inspiration of artificial systems, it is called *stigmergy* (name introduced by P.P.Grassé in the 50's). Following Wilson (1971), the hypothesis of stigmergy is that

> "(...) it is the work already accomplished, rather than direct communication among nest mates, that induces the insects to perform additional labor."

First of all, it is clear why stigmergy it is interesting for RDAI: The problem of communication between agents simply vanishes. So what remains is to build a working individual operating in a common environment with the other individuals. There is a plethora of phenomena which can be explained by means of stigmergic mechanisms, such as nest building in termites (Deneubourg, 1977), collective decision making (shortest path between the nest and some other place and selection of the best food source, in section 3.1.1 there is a detailed explanation of a model of stigmergy in collective decision, see Millonas, 1992, for further details) and collective clustering (see section 3.1.1 for details) to cite a few. Artificial systems built according to an underlying stigmergic mechanism are simply the majority of the systems built up to now. Examples are the above detailed works on the TSP or the pushing robots and the mechanism of implicit cooperation, used mainly with *real* robots for obvious cost reasons (see Mataric, 1994; Deneubourg *et al.*, 1991, 1992; Goss & Deneubourg, 1992; Kube & Zhang, 1993; Sugawara & Sano, 1997).

Stigmergy is also related with the slaving principle, since the work done by individuals acts as a sort of "slow" variable slaving, and governing, the individual behaviour. In the example of clustering, the objects forming small clusters induce the individuals to put down objects near the clusters, creating a positive feedback that ends with one or two large clusters (section 3.1.1, Deneubourg *et al.*, 1991); in the example of the TSP and also that of collective decision, the "trail" (section 1.1, Colorni *et al.*, 1992) and the morphogen (section 3.1.1, Millonas, 1992) have "slow" dynamics due to changes in concentrations induced by individual activity, though they also act as a field governing individual behaviour. We see here a *circular causality*, one of the main features of complex systems (figure 1.4).

1.5 On collective computation

The world of ants is so huge that it is quite likely that the phenomena as yet hidden to us will be far more surprising than those facts already known. Undoubtedly RDAI will benefit from these findings and new design procedures for artificial collective systems will see the light in a near future. There are a plethora of mechanisms by means of which ant colonies perform their striking collective behaviors, and the one most exploited up to now, stigmergy, is just one of them. Thus, it is an open question which unexplored (from a RDAI point of view) mechanisms may be useful to design artificial systems. In this thesis we will explore a recently discovered temporal pattern in ant colonies: Self-synchronization (Cole, 1991a, 1991b; Franks & Bryant, 1987; Franks *et al.*, 1990). It has been shown that self-synchronization underlies

some quite interesting phenomena to RDAI practicioners, such as task allocation (Robinson, 1992; Hatcher *et al.* , 1992), pointing to an interesting relation between a self-organized temporal behavior and some useful activity. See chapter 2; section 2.1 introduces a detailed account of the biological facts and the rest of the chapter introduces and studies a mathematical model called *Fluid Neural Network.*

Another interesting problem is that of the capabilities and limitations of collective systems *in general.* Are collective systems restricted to a certain set of possible tasks? Is there some threshold in individual simplicity beyond which any useful task is impossible? or, on the contrary, are non-simple individuals unable to display emergent functional behavior? If we do not restrict the (not formally defined) class of "collective systems" some answers are already at hand, since collective systems such as cellular automata and neural networks are able of universal computation, by means of Turing machines simulation (Garzon, 1995). Collectives of concurrent simple agents can also simulate any Turing machine, as was shown with the process algebra CCS in Milner (1990). Furthermore, it is widely believed that some problems are "inherently sequential" (Balcázar *et al.*, vol. 2, 1990), so it is clear that not in all problems collective systems are more efficient than hierarchical and sequential systems. However, this is not our goal. We want to restrict the class of "collective systems" to *real* systems, that is, either natural or artificial physical collective systems displaying emergent functionality, and we want to study these systems *with computational means.* In other words, our departure point is the real system (composed of ants or robots), which is properly modeled with a dynamical system. Our hypothesis is that *computational information may be obtained from a (computational) study of a (good) model of the real system.* With this goal in mind, we will explore some ways of relating dynamical systems with computational notions in chapters 3 and 4. In chapter 3 we introduce the problem at length and we explore the possibility of computing with Fluid Neural Networks. Some drawbacks appear and we resort in chapter 4 to the new field of *Computational Mechanics* to relate dynamical systems and computation, in order to formalise an interesting problem relating individual and collective capabilities.

These are the problems dealt with in this thesis and what has been achieved will be discussed in chapter 5 (overview and prospects). The more classical (from a physicist point of view) approach to the study of Fluid Neural Networks is what has allowed us to go that far in the analytical study of the system, though some work remains to be done to get a completely satisfactory (mean field) theory of Fluid Neural Networks. With respect to the rest of the thesis, we have decided on a more innovative approach to interesting problems currently lacking an adequate

treatment, as we see it. Our approach, based on Crutchfield's Computational Mechanics, tries to open a door to an immense landscape where nature is, for the first time, examined *systematically* with computational tools. However, we are merely starting to push at that door.

Chapter 2

Noise-Induced Phenomena and Fluid Neural Networks

2.1 Self-synchronized behaviour in ant colonies

Ants, refuting popular fables, do not work untiringly all day long. As Sudd (1967) pointed out, the proportion of time spent in resting can be high, and the study of Herbers (1983) on acts performed by ants of *Leptothorax longispinosus* and *L. ambiguus* species reveals that "(...) ants spent two-thirds of their time apparently doing nothing at all". These behavioural patterns are by no means exceptional, Franks & Bryant (1987) found them in *L. acervorum* and Cole (1986) noted that "(...) Time spent quiescent occupies a large fraction of the total time of an ant (on average 55%)" while studying ants of the species *L. allardycei*. A more refined study by Franks *et al.* (1990) measured even a 72% of time spent resting for workers inside nests of *L. acervorum*. However, not only patterns of alternate activation have been found in individual ants, these patterns also appear *in whole colonies, showing synchronized patterns of activity*: surprisingly, Franks & Bryant (1987), by means of video-recording techniques, were able to get a long enough time series of the activity in whole colonies of *L. acervorum* to show, using spectral analysis, that activity was roughly periodic, with periods between 15 and 30 minutes. This synchronized behaviour has also been found in *L. longispinosus*, *L. ambiguus*, *L. curvispinosus*, *L. allardycei* and *L. muscorum* (see Miramontes 1992, chap.2), also in other species such as *Pseudomyrmex elongatus*, *P. pallidus*, *Tapinoma littorale*, *Zacryptocerus varians* and *Crematogaster ashmeadi* (Cole & Cheshire, 1996). Activity patterns are not just synchronized, but *self*-synchronized: no external signal has been found experimentally as a possible cause of colony synchronization (Cole, 1991a).

In order to situate the discussion in its proper context let us summarize some biological features of the genus *Leptothorax*. P.A. Latreille introduced the term

Leptothorax almost 200 years ago, though the taxonomical status of this genus is not clear: we can find strange properties in some of the nearly 350 named taxa of the *Leptothorax*, such as members of the same species, *L. muscorum*, having different chromosomal numbers (Miramontes, 1992). The distribution of *Leptothorax* is worldwide: from Alaska (*L. acervorum*) to Florida (*L. allardycei*). *Leptothorax* colonies have around 100 members, though this number can oscillate between 10 and 500 individuals, it depends essentially on the age of the colony: Tofts *et al.* (1992) cite *L. acervorum* colonies with a number of individuals among 12 and 216 members, Cole (1991a) finds a range of 20-120 individuals in *L. allardycei* colonies and Franks *et al.* (1992) find that colonies of *L. unifasciatus* have among 60 and 184 members (let us refer, in passing, to the striking contrast among ant species with respect to the number of individuals, just mentioning the more than 300 million members of *Formica yessenis* colonies). A colony of *Leptothoracine* ants is typically composed of females (males are just for mating and fertilizing eggs), whose majority are workers engaged in daily colony activities. The absence of a queen is not a rare fact: Alloway *et al.* (1982) reported that a 29.7% of 1522 observed colonies of *L. ambiguus* were queenless, a 36.6% of 488 colonies of *L. curvispinosus* were also queenless and the same was observed in a 37% of 640 *L. longispinosus* colonies. Finally, ants of the *Leptothorax* genus are monomorphic (there is no physical difference among ants belonging to different castes, therefore caste division is correlated with age, see Hölldobler & Wilson, 1990) using a short-distance mode of communication, mainly body to body contact, though, as other ant species, *Leptothorax* ants have well developed glands capable of producing chemicals for communication purposes (Herbers, 1983). Self-synchronization, as we already said, has been found in diverse *Leptothorax* species, but the work of Cole (1991a) is particularly interesting, since he studies also individual activation dynamics, obtaining data quite relevant for designing individual-based mathematical models of oscillatory behaviour (see section 2.2). The method used in Cole (1991a) to collect data from individuals and colonies was based on recording images of whole colonies of *L. allardycei* every 30 seconds, measuring activity levels by taking the pixel differences between two successive images (see figure 2.1, top plot). The analysis was performed by means of periodograms (see figure 2.1, middle plot) with peaks in the Fourier components around a period of 27 min. per cycle, and autocorrelation functions (see figure 2.1, bottom plot), whose sinusoidal nature indicates clearly that the time series is periodic, with a mean period of approximately 26 min. (see Cole, 1991a for details). The same sort of measures performed over a single isolated ant made evident spontaneous activation and quiescence during long periods of time but no periodic activity. Fur-

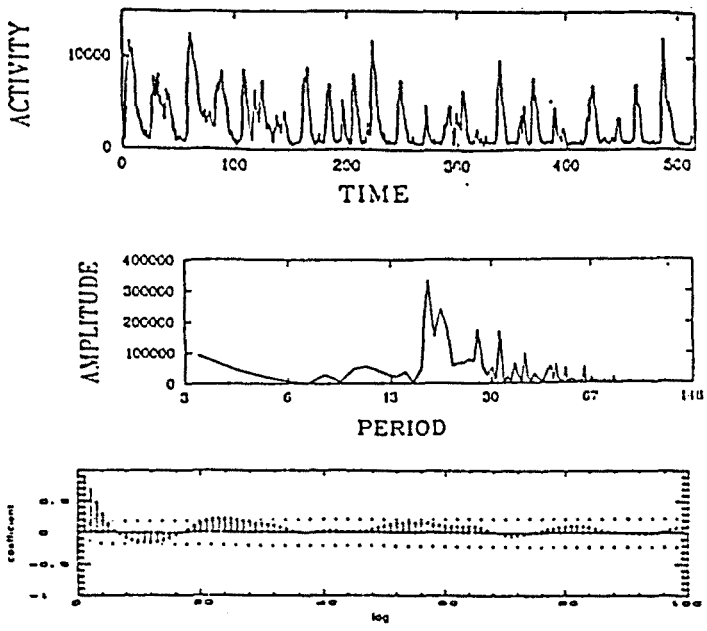


Figure 2.1: On top: activity record for one colony of *L. allardycei*, time measured in 30 sec. intervals and activity measured by pixel differences, see text. Middle: Periodogram of the colony's activity record: squared amplitude of Fourier components *vs* the natural logarithm of the frequency. Bottom: Autocorrelation function of activity record, see text (after Cole, 1991a)

thermore, Cole (1991b) was able to show evidence of chaotic activity in single ants. So, one important conclusion is already at hand, that is, *self-synchronization is a collective property*, since individual patterns of activation are not periodic. Finally, Cole (1991a) discusses the adaptive significance of short-term activity cycles arguing that it is unlikely that these cycles contribute to the efficiency of the colony. They are "(...) the inevitable outcome of interactions within social groups".

However, at least two functional behaviours in ant colonies have been related to self-synchronized activity: task allocation (Robinson, 1992) and mutual exclusion (Hatcher *et al.*, 1992). Task allocation in ant colonies is a extremely fascinating problem, making evident, perhaps in its very essence, the collective performance of insect societies: each ant in a colony seems to know exactly what to do in order to fulfill *global* colony needs (see sections 2.5 and 3.1.1. for a more detailed description of task allocation in ant colonies). It seems that self-synchronized behaviour provides a mechanism for information propagation:

> "Sampling behavior that involves social interactions may be facilitated by synchronous bursts of worker activity, which have been observed in

> ant colonies (...). The decision of which task to perform would be based on the integration of acquired information, coupled with behavioral biases associated with worker, caste, physiological status and prior experience." (Robinson, 1992, p. 652)

Mutual exclusion in *L. acervorum* colonies has been proposed as a mechanism for effective exchange of information on task allocation (Hatcher *et al.*, 1992). Inside nests of *L. acervorum* nurse workers interact in order to determine which items of brood require attention, with the constraint that no more than a few nurses can tend a brood item (spatial arrangement of brood limits the number of workers that can access brood simultaneously). Assume that nurse workers choose at random which brood item to tend. In this situation, some brood items may be ignored during a too long period of time, long enough to endanger their survival. Assume that the probability of tending a certain brood item is $1/B$ (there are B brood items and A nurse workers), then the probability that no nurse worker tends that brood item is $(1 - 1/B)^A$. So, the proportion P_{random} of brood tended in any period of time will be

$$P_{random} = 1 - \left(1 - \frac{1}{B}\right)^A \tag{2.1}$$

If we had synchronized activity, the situation would be quite different. In this case, all nurse try to tend some brood item, what causes a even distribution due to the spatial access constraint. The proportion $P_{exclusion}$ of brood tended in a period during which each ant is active once (an accounting period) is

$$\begin{aligned} P_{exclusion} &= A/B \qquad \text{for } A < B \\ P_{exclusion} &= 1 \qquad\quad \text{for } A \geq B \end{aligned} \tag{2.2}$$

Figure 2.2 makes clear the superior efficiency of mutual exclusion mechanism, that is, self-synchronization plus spatial restrictions in brood access. So, self-synchronization is the mechanism behind some interesting functional behaviours in ant colonies. The above mentioned examples also clarify why we think it is worth to study in depth self-synchronization from a RDAI point of view. After self-synchronization had been well understood, it should be as fruitful as has been the understanding of stigmergic mechanisms (see section 1.4) to the design of algorithms for collective behaviour.

Two main observations/conclusions emerge from these experiments.

- Randomness is present at the individual level, acting as a source of noise.

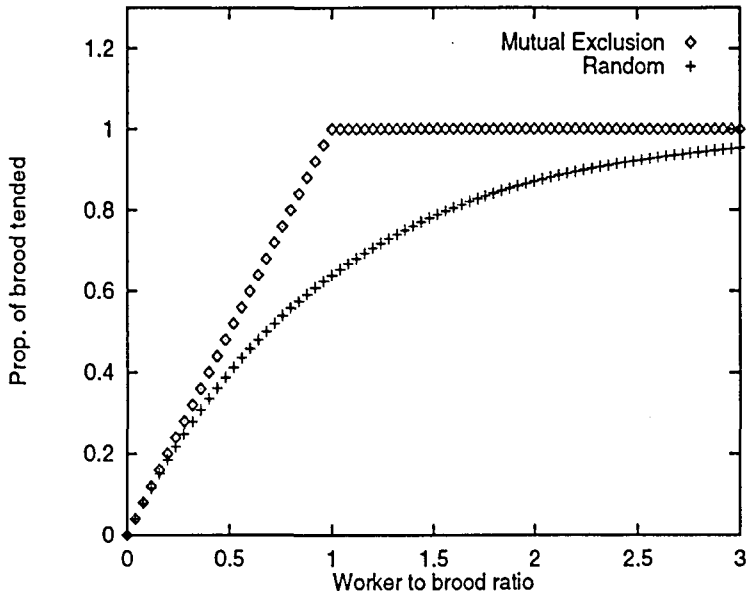


Figure 2.2: Proportion of brood tended with respect to worker-to-brood ratio for mutual exclusion and random models. It is obvious the better efficiency of mutual exclusion (after Hatcher et al., 1992)

- Collective ordered dynamics emerge from the microscopic chaos. The observed oscillations are present in different species and seem to be an intrinsic relevant feature of ant colonies.

These observations lead to some relevant questions, some of which are answered (at least under some approximation) in this chapter:

- What is the link between the stochastic nature of individuals and the outcome of their interactions?
- What kind of collective behavior is shown by these groups of simple agents? More precisely, what kind of collective (physical) phenomenon is at work?
- Is a transition-like phenomena responsible of the collective dynamics? How can such type of phenomenon be characterized through an order parameter?
- If the self-organized oscillations underlie a computational process of some kind (task allocation, for example), why are complex fluctuations used by ant colonies, instead of constant levels of activity?

As we will see, the random-like behavior of single ant dynamics could be used as a source of noise inside a nonlinear system where fluctuations are amplified. As discussed in the introduction, such kind of interactions between noise and nonlinearity

can be the very source of complex behavior in many real physical systems. A wide range of problems have been explored in the physics literature, ranging from fluid dynamics to condensed matter. Here we will explore a phenomenon which deals with how computation may emerge from noisy interactions.

2.2 Fluid Neural Networks

2.2.1 Definition

The relation of brains with ant colonies, strange as it seems, is a profound one. It has been noted in the past:

> "We must also add that electronic engineers have now constructed circuits in which the different parts are joined to each other by as many connections as possible. If the connections are both numerous and random the whole network then has certain properties which remaind one of the brain (...) I well know that an ant nest is not a brain, but this does not mean that the basic principles of their organization are not similar, or that the study of the ant colonies cannot teach us something about the brain and vice versa. That would be a most unexpected result of the study of myrmecology ..." (Rémy Chauvin, cited in Solé *et al.*, 1993b)

and also

> *Achilles:* Familiar to me? what do you mean? I have never looked at an ant colony on anything but the ant level.
>
> *Anteater:* Maybe not, but ant colonies are no different from brains in many respects...
>
> (Hofstadter, 1979)

Both, brains and ant colonies, have some common properties, such as to depend on collective properties to achieve full functionality while being composed of relatively simple elements. Besides, the behaviour of these simple elements gives no information on the global behavior of the system: new emergent phenomena arise due to interactions among such components.

An explicit comparison between both systems (ant systems and generic biological neural nets) is shown in the following table, where several characteristics (both

dynamical and structural) are considered. In both systems, typical situations are assumed.

	Ant Colonies	Neural Networks
Number of units	high	high
Robustness	high	high
Connectivity	local	local
Memory	short-term	short/long-term
Stability of individual connections	weak	high
Global spatial pattern of activity	trails	brain waves
Complex dynamics (1/f)	Observed	Common

It is not surprising that the main differences arise from connectivity: Direct contact among individual ants is a transient phenomenon. On the contrary, synaptic connections among neurons has a characteristic lifetime which is similar to the lifetime of neurons. A direct consequence is that memory in ant colonies will be typically short. This is partially compensated (particularly in large colonies) by the use of chemicals, which can create spatial structures that clearly involve (long term) memory effects.

Our theoretical approach to modelling self-synchronization in ant colonies will be based on the above mentioned similarities between brains and ant colonies (see Gordon *et al.*, 1992 and Solé *et al.*, 1993a and 1993b) introducing the *Fluid Neural Network* (FNN), a term first coined by R.V. Solé in 1993. In FNN the standard approach of neural networks is used (Amit, 1989), but a new set of rules defining local movement and individual activation are also introduced. A set of N automata or "neuron-ants" is used. The state of each automaton (say the i-th one) is described through a continuous state variable $S_i(t) \in \mathbf{R}$, at each time step $t \in \mathbf{N}$. Each element can move on a $L \times L$ two-dimensional lattice (figure 2.3). A set of rules is defined:

(a) Neural network structure: Interactions are described as in neural networks by means of a sigmoidal function $\Phi(x)$. If $S_j(t)$ is a given automaton (the spatial dependence is omitted for simplicity), the new states are updated following:

$$S_i(t+\Delta t) = \Phi\left[\; g\left\{ J_{ii}S_i(t) + \sum_{i\neq j\in B(i)} J_{ij}S_j(t) \;-\Theta_i \right\}\right] \qquad (2.3)$$

where $B(i)$ are the nearest automata, located in the neighborhood defined by the eight nearest lattice sites, and $J_{ii} \neq 0$. For simplicity we use the threshold $\Theta_i = 0$, and we take $\Phi(z) = \tanh(gz)$ with g being a gain parameter.

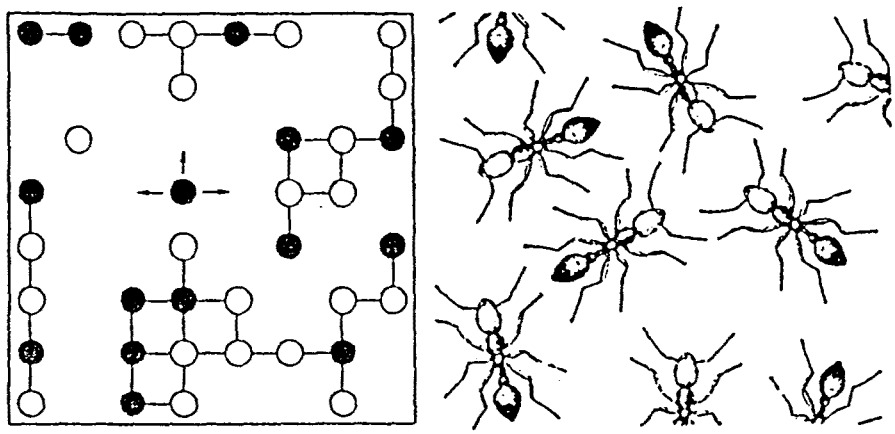


Figure 2.3: Our ants will be formal neurons, "neuron-ants", capable of movement on a $L \times L$ lattice. Only active individuals (black circles) will move, if possible. Inactive elements (white circles) may activate spontaneously or by interaction with active individuals.

(b) Spontaneous activation: We have seen above that one of the properties Cole (1991a) observed in isolated ants was spontaneous activation. In FNNs this has been included in the following way: each automaton can be either *active* or *inactive* and, if active, it moves randomly to one of the eight nearest cells (if no space is available, no movement takes place). In our model a given automaton will be active if $S_i(t) > \theta_{act}$ and inactive otherwise. Once an automaton becomes inactive, it can return to the active state (with an *spontaneous activity level* S_a) with some probability p_a.

(c) Coupling matrix: the coupling matrix **J** is not fixed. Connections are local and changing over time as a consequence of movement. They are also state-dependent i.e. J_{ij} will be a simple function of the states of the actually interacting pair (i,j) of automata, i.e. $J_{ij} = f(a_i^t, a_j^t)$, where $a_i^t = \Theta[S_i(t) - \theta_{act}]$ ($\Theta[x]$ is the Heaviside step function). In our case, where two basic states are defined, i.e. "active" and "inactive", the connectivity matrix reduces to the following 2×2 table:

$$\Lambda = \begin{pmatrix} \lambda_{11} & \lambda_{10} \\ \lambda_{01} & \lambda_{00} \end{pmatrix} \tag{2.4}$$

At a given time step, the interaction J_{ij} between the i-th and the j-th elements is equal to $\lambda_{a_i^t a_j^t} \in \Lambda$ by depending on the activity states of the given elements. More precisely, J_{ij} will be equal to: λ_{11} when both ants are active, to $\lambda_{10}, \lambda_{01}$ when one is active and the other inactive and to λ_{00} if both automata are inactive. In this thesis we take for simplicity $\lambda_{a_i^t a_j^t} = 1$. Our choice is based in the observation of ant colonies. Self-interaction and positive feed-back (with local excitability) play an

essential role in colony dynamics. The consistency of our choice is fully supported by experimental studies of ant colonies (Cole & Cheshire, 1996).

A technical point is what sort of boundary conditions we use. Though not biologically plausible, we use *periodic* boundary conditions, because we want to verify that ordered temporal behaviour is a consequence of the *intrinsic* dynamics of the system, not a boundary effect. Nevertheless, this turns out to be irrelevant because of the same phenomenology is observed with diverse boundary conditions (periodic and zero-flux).

2.2.2 Analysis of individual behaviour

First of all we have to prove that there is no periodic behaviour in individual automata, in order to assure that collective oscillations are really collective. Our individual, as defined in the previous section, change state $S(t)$ according to

$$S(t+\Delta t) = \tanh(g\lambda_{a^t a^t} S(t)) \tag{2.5}$$

where $a^t = \Theta\left[S(t) - \theta_{act}\right]$ and $\lambda_{a^t a^t} = 1$. If inactive, i.e. $a^t = 0$, it may activate spontaneously with probability p_a and spontaneous activity S_a. Our individual will be quite similar to a classic Poisson process, except for the time τ_{ind} during which the individual is active and no spontaneous activations can occur. This fact allows us to compute the probability of having n spontaneous activations in a time interval of length T, $P_n(T, \tau_{ind})$, provided T is long enough and Δt is very small. If $\tau_{ind} = 0$ we have a Poisson process and the above mentioned probability is a well-known result of probability theory (Ricciardi, 1977)

$$P_n(T,0) = \frac{(p_a T)^n \, e^{-p_a T}}{n!} \tag{2.6}$$

When $\tau_{ind} > 0$, the calculation of $P_n(T, \tau_{ind})$ is more involved. We will follow Ricciardi (1977) and exploit the analogy between our randomly activated "neuron-ant" and a model neuron with absolute refractoriness τ_{ind} subject to a Poisson sequence of excitatory zero-width point inputs with arrival rate p_a. To calculate $P_n(T, \tau_{ind})$ we replace each zero width pulse occurring at, say, t_i by a pulse of duration τ_{ind} beginning at t_i. The effect of refractoriness will be to prevent any spontaneous activation in the time interval $(t_i, t_i + \tau_{ind})$.

First let us consider the case of $n < T/\tau_{ind}$. Let S be the set of all n-tuples of spontaneous activations, occurring at $t_1, t_2, \ldots, t_n$ in the interval $(0, T)$, and S_1 and S_2 two subsets of S such that $S_1 \cup S_2 = S$ and $S_1 \cap S_2 = \emptyset$. S_1 is the set of

n-tuples verifying $t_n + \tau_{ind} < T$, and S_2, obviously, is such that $t_n + \tau_{ind} \geq T$. Let the probability of a particular sequence of spontaneous activations in S_1 at $t_1, t_2, \ldots, t_n$ $(t_i + \tau_{ind} \leq t_{i+1})$ be $d^n P_n^{(1)}(T, \tau_{ind})$. This probability is simply the probability of having n pulses at t_i times the probability of no pulses in the interval $(t_i + \tau_{ind}, t_{i+1})$. From $P_n(T, 0)$ (eq. 2.6) and the known fact that in a Poisson process with arrival rate p_a the probability of one pulse in a time interval $(t, t + dt)$ is approximately $p_a dt$ one can compute

$$d^n P_n^{(1)}(T, \tau_{ind}) = p_a^n e^{-p_a(T - n\tau_{ind})} dt_1 dt_2 \ldots dt_n \tag{2.7}$$

A similar reasoning yields, for n-tuples in S_2

$$d^n P_n^{(2)}(T, \tau_{ind}) = p_a^n e^{p_a(n-1)\tau_{ind}} e^{-p_a t_n} dt_1 dt_2 \ldots dt_n \tag{2.8}$$

Because S_1 and S_2 are a disjoint partition of S, to get the desired result $P_n(T, \tau_{ind})$ it will be enough to add $d^n P_n^{(1)}(T, \tau_{ind})$ for all the n-tuples in S_1 and $d^n P_n^{(2)}(T, \tau_{ind})$ for all the n-tuples in S_2, that is

$$P_n(T, \tau_{ind}) = \int_{S_1} d^n P_n^{(1)}(T, \tau_{ind}) + \int_{S_2} d^n P_n^{(2)}(T, \tau_{ind}) \tag{2.9}$$

The first integral can be performed considering eq. 2.7 and the possibilities to locate t_i (from a n-tuple of S_1) in the interval $(0, T)$:

$$\begin{aligned}
0 \leq t_1 &\leq T - n\tau_{ind} \\
t_1 + \tau_{ind} \leq t_2 &\leq T - (n-1)\tau_{ind} \\
&\vdots \\
t_{n-1} + \tau_{ind} \leq t_n &\leq T - \tau_{ind}
\end{aligned}$$

yielding

$$\int_{S_1} d^n P_n^{(1)}(T, \tau_{ind}) = \frac{p_a^n (T - n\tau_{ind})^n}{n!} e^{-p_a(T - n\tau_{ind})} \tag{2.10}$$

similarly we obtain the second integral

$$\begin{aligned}
\int_{S_2} d^n P_n^{(2)}(T, \tau_{ind}) = e^{-p_a[T - n\tau_{ind}]} \sum_{k=0}^{n-1} \frac{p_a^k \left[T - n\tau_{ind}\right]^k}{k!} - \\
- e^{-p_a[T - (n-1)\tau_{ind}]} \sum_{k=0}^{n-1} \frac{p_a^k \left[T - (n-1)\tau_{ind}\right]^k}{k!}
\end{aligned} \tag{2.11}$$

Using eq. 2.9 we have got $P_n(T, \tau_{ind})$ when $n < T/\tau_{ind}$. In the case $T/\tau_{ind} \leq n \leq T/\tau_{ind} + 1$ we can find $P_n(T, \tau_{ind})$ using the same reasoning above mentioned though it is much more tedious (it will not be detailed here, see Ricciardi, 1977). At last, we obtain, for all n and T

$$P_n(T, \tau_{ind}) = \Theta(T - (n-1)\tau_{ind}) \Bigg\{ 1 - e^{-p_a[T-(n-1)\tau_{ind}]} \times \\ \times \sum_{k=0}^{n-1} \frac{p_a^k \left[T - (n-1)\tau_{ind}\right]^k}{k!} - \Theta(T - n\tau_{ind}) \times \\ \times \left[1 - e^{-p_a[T-n\tau_{ind}]} \sum_{k=0}^{n} \frac{p_a^k \left[T - n\tau_{ind}\right]^k}{k!} \right] \Bigg\} \tag{2.12}$$

Finally we can approximate τ_{ind} by linearising $\tanh(x)$

$$S(t + \Delta t) \simeq g\lambda_{a^t a^t} S(t)$$

and defining activity as starting from the spontaneous activity level $S(0) = S_a$. The state after τ iterations will be

$$S(\tau) \simeq \left(g\lambda_{a^t a^t}\right)^\tau S_a$$

from which we can easily obtain

$$\tau_{ind} \simeq \frac{\log_{10}(\theta_{act}/S_a)}{\log_{10}(g\lambda_{a^t a^t})} \tag{2.13}$$

So then, the probabilistic nature of our individuals has been proved. Whether there is collective ordered behaviour in the system as a whole, it has to be caused by interactions among individuals and not by the individuals themselves. If we observe global spatial or temporal structures, they will be truly *emergent* properties.

2.2.3 Collective behaviour of FNNs: Oscillations

The collective behaviour we want to measure in FNNs is the mean activity of the system. We have defined above an activity for each individual $S_i(t)$, $a_i^t = \Theta\left[S_i(t) - \theta_{act}\right]$, so the mean activity at time t will be

$$\rho_t^+ = \frac{1}{N} \sum_{j=1}^{N} a_j^t = \frac{1}{N} \sum_{j=1}^{N} \Theta\left[S_i(t) - \theta_{act}\right] \tag{2.14}$$

where $\rho_t^+ \in [0, 1]$. We define also the total density of automata as $\rho = N/L^2$. We can see in figure 2.4 a sample of the temporal behaviour of ρ_t^+ for different

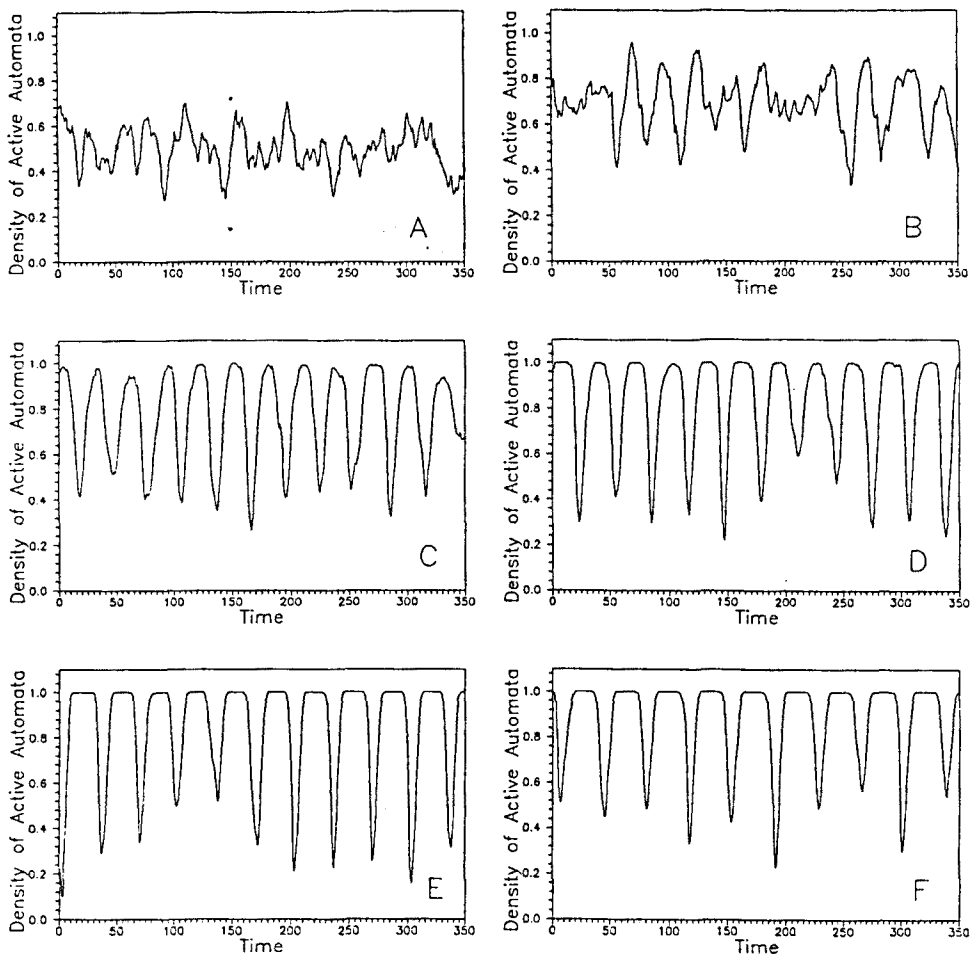


Figure 2.4: Temporal behavior of ρ_t^+, with parameters $L = 50$, $S_a = 0.1$, $g = 0.1$, $\theta_{act} = 10^{-16}$, $p_a = 0.01$ and (A) $\rho = 0.10$; (B) $\rho = 0.15$; (C) $\rho = 0.20$; (D) $\rho = 0.25$; (E) $\rho = 0.30$; (F) $\rho = 0.35$;

densities ρ. At low ρ we have a very disordered pattern, which becomes more and more ordered as ρ is increased. So, a clear transition from random to ordered, *self-synchronized* temporal behaviour is made evident, at least visually. Solé *et al.* (1993a) calculated the Fourier spectrum for time series as those shown in figure 2.4. A well defined maximum in the amplitude appears with growing ρ. Note that despite our focusing on a quantity that apparently does not include the spatial dependence, this dependence is partially responsible of the dynamics of ρ_t^+.

The time evolution of ρ_t^+ is a stochastic process because of the spontaneous activation of inactive automata. The process ρ_t^+ is non-Markovian due to the existence of the time interval in which the automaton is active, interacting with other automata. In the absence of interactions this time interval is the above mentioned τ_{ind}: the number of active automata at time t will be those active at $t-1$ plus those

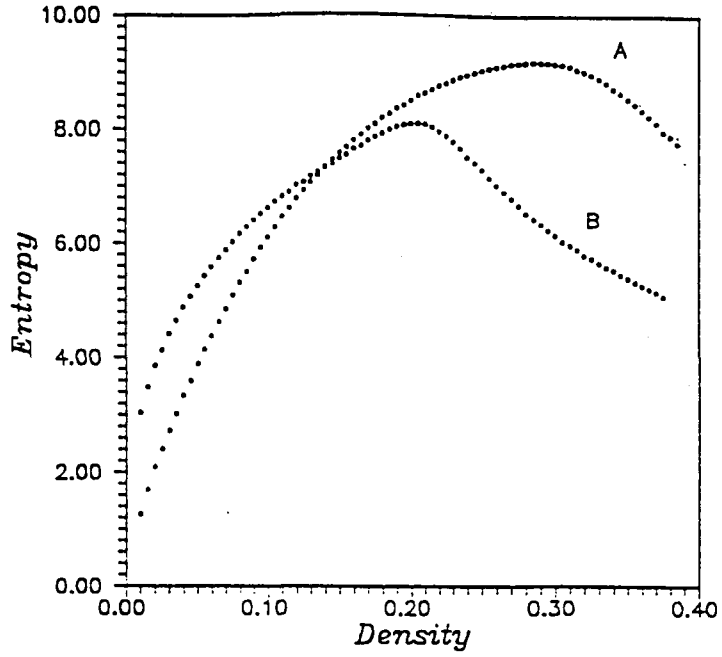


Figure 2.5: SK entropy vs. Density, with parameters $L = 50$, $S_a = 0.1$, $g = 0.1$, $\theta_{act} = 10^{-16}$ and (A) $p_a = 0.001$; (B) $p_a = 0.01$. Here $T = 2 \times 10^4$ after 2×10^3 transients were discarded.

inactive at $t - 1$ that spontaneously become active, minus those that were active at $t - \tau_{ind}$; however, if we consider the interactions between automata, this time interval is altered by the patterns of interaction.

Solé & Miramontes (1995) showed that a first approach to a quantitative characterization of the transition in the dynamical behavior of the FNN was to measure the *Shannon-Kolmogorov entropy* of the following probability distribution $\{p(j,\rho)\}$: let T_j the number of time steps where exactly j elements were active ($j = 0, 1, ..., N$); now, if T is the total number of time steps, the relative frequency of j simultaneously active elements will be $p(j,\rho) = T_j/T$. The SK entropy is then defined as:

$$S(\rho) = -\sum_{i=0}^{N} p(i,\rho)\ \log_2\ p(i,\rho) \tag{2.15}$$

It is expected that SK entropy will provide us with a measure of the diversity of macroscopic states. $S(\rho)$ has an upper and a lower bound: $S(\rho) \in [0, \log_2(N+1)]$. At low densities, $S(\rho)$ will grow until the onset of collective oscillations, where it will decrease due to the highly correlated temporal patterns (see figure 2.5). So, the transition occurs at a *critical density* ρ_c such that $S(\rho_c)$ is maximum.

Using SK entropy, Solé & Miramontes (1995) performed an extensive study of the (ρ, g)-space. They found four dynamical domains: random behaviour, chaotic

attractors, steady states and periodic oscillations. For high g values the automata collapses to a steady state where all individuals are active. Lower g results in coherent oscillations or disordered behaviour (Solé & Miramontes, 1995, called it "chaotic") depending on the density ρ. These domains are separated by a maximum in SK entropy. Finally, there is a region $g < g_c$ where no transition arises and SK entropy always grows with ρ as the logarithm of the number of objects $S(\rho) \simeq \log_2(N)$, all states tend to be equally represented hence it is called "random".

Another measure often used when dealing with critical phenomena is the *transient length*. Now, Solé & Miramontes (1995) calculated the probability distribution $\{P_s(r)\}$ where r means "active" (1) or "inactive" (0) over a long number of time steps ($\simeq 10^4$), in order to make sure its stationarity. They also calculated a time dependent set of probabilities $\{P(r;t)\}$ in the same way that they calculated $\{P_s(r)\}$ but only up to time t, in such a way that $\{P(r;t)\}$ tends asymptotically to $\{P_s(r)\}$. To compare both probability distributions they used the *information gain* or *Kullback information*

$$K(P_s, P; t) = \sum_{r=0}^{1} P(r;t) \log_2 \left[\frac{P(r;t)}{P_s(r)} \right] \tag{2.16}$$

$K(P_s, P; t)$ has the important property $K(P_s, P; t) \geq 0$ where equality holds if and only if $P(r;t) = P_s(r)$ for all r (this property is easily proved from the general inequality $\ln(x) \geq 1 - x^{-1}$ which holds for all $x > 0$). The transient length l was defined as the first time step such that

$$K(P_s, P; l) \leq \epsilon \tag{2.17}$$

where $\epsilon = 0.0025$. This process is repeated over a number of samples and averaged. As we can see in figure 2.6, l reaches its maximum also when collective oscillations appear.

Thus, a transition from disordered to oscillatory behaviour in FNNs has been characterized numerically as the maxima of both SK entropy and transient length. Now, how good are FNNs to model the *real* system? In principle the same phenomenology has been observed. However, in figure 2.5 we observe the onset of oscillations for a critical density $\rho_c \simeq 0.2$, *the same density observed in real ant colonies* (Franks *et al.*, 1992). This is clearly a point in favour of FNNs as models of self-synchronized behaviour in *Leptothoracine* ant colonies, besides providing the first evidence of real systems living at the edge of a phase transition (Solé & Miramontes, 1995). On the other hand, despite having identified a transition, which kind of transition is it? Non-equilibrium systems (and FNNs are obviously not in

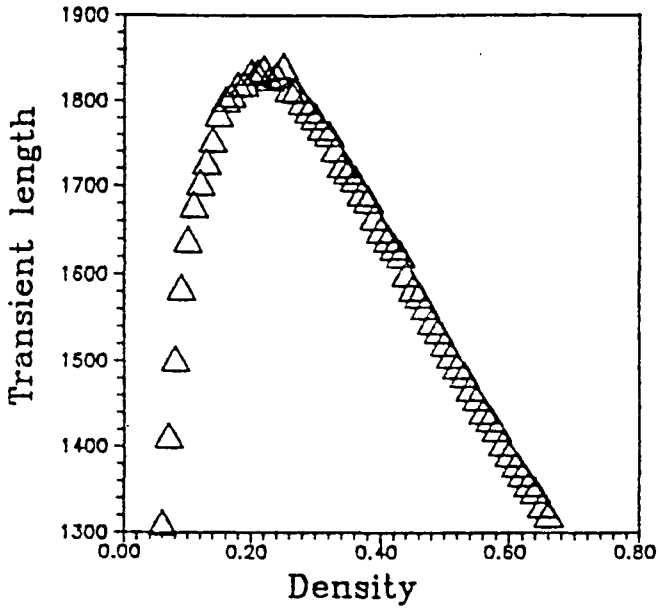


Figure 2.6: Transient length l (using 100 samples for averaging). It becomes maximum at the transition point (after Solé & Miramontes, 1995)

equilibrium) may display different types of transitions (see the introductory chapter of Horsthemke & Lefever, 1984, for an excellent survey of non-equilibrium phase transitions). This is the subject of next section.

2.3 Phase Transitions in FNNs

We have surveyed the study on FNNs by Solé & Miramontes (1995) concerning a quantitative characterization of disorder-order transitions, summing up their work on the g dependence of the critical density ρ_c. We will see in this section that it will be more advantageous to focus on the relation between ρ_c and the probability of spontaneous activation p_a, since it will allow us not only to characterize quantitatively the transition, but also to relate that transition to the theory of non-equilibrium phase transitions.

2.3.1 Noise Induced Transitions: A Brief Introduction

It is nowadays well known that systems out of equilibrium may display phase transitions quite similar to equilibrium phase transitions and phase transitions without counterpart in equilibrium phenomena (Haken, 1977). An example of the latter are the *noise-induced transitions* (Horsthemke & Lefever, 1984), transitions linked to the response of some *nonlinear* systems to *multiplicative* (state-dependent) ex-

ternal noise. The general scenario of noise-induced phase transitions starts from a stochastic differential equation

$$\frac{dx}{dt} = \phi_\mu(x) = h(x) + \lambda_t g(x) \tag{2.18}$$

where λ_t stands for a stochastic component which depends on the state of the environment. λ_t takes into account the stochastic nature of such environment by splitting in two parts

$$\lambda_t = \lambda + \sigma\xi_t$$

where the external white noise ξ_t has zero mean and intensity σ^2. We can see that this description is rather different from the linear noise term presented in the introductory chapter (section 1.2). In that case the noise term was additive, and so then a stochastic description of the dynamical system lead to probability distributions centered around the deterministic steady states. Here, however, noise is multiplicative and so a communication between fluctuations and the macroscopic behavior becomes still more relevant. Now fluctuations are not only amplified, but they are also largely responsible of how the system is organized far beyond the deterministic description. These processes display transitions that could modify the bifurcation diagrams in a much more profound way than just by a ahift in parameter space. What is more important in the context of pattern recognition and information processing: do nonlinear systems, coupled to a fluctuating environment, always adjust their macroscopic behavior to the average properties of the environment, or can one find situations in which the system responds, in a certain *more active* way to the randomness of the environment, displaying, for instance, behavior forbidden under deterministic external conditions?

This question has a positive answer (Horsthemke & Lefever, 1984) and a well defined theoretical framework has been developed. In a close analogy with the additive noise case discused in chapter 1, a Fokker-Planck analysis can be performed. For our problem, described by 2.18, the stationary solution for the probability distribution $P_s(x)$ now reads

$$P_s(x) = \frac{N}{g^2(x)} \exp\left[\frac{2}{\sigma^2}\int^x \frac{f(u)}{g^2(u)}du\right]$$

As an example, let us consider the genetic model (Horsthemke & Lefever, 1984, sect. 6.5)

$$\frac{dx}{dt} = 0.5 - x + \lambda x(1-x) \tag{2.19}$$

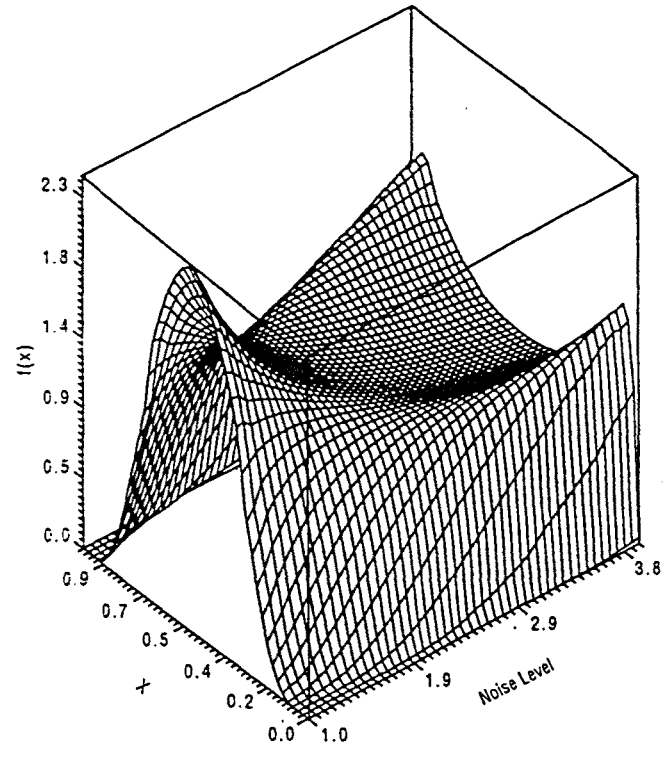


Figure 2.7: Noise-induced phase transition. Here the deterministic dynamical system given by eq. 2.19 is perturbed by means of a multiplicative noise term. The stationary probability density (as given by eq. 2.20) is shown. As the noise level is increased, the single maximum is replaced by two new extrema.

with $x \in [0,1]$. In a constant environment (λ constant) no instability occurs, there is no transition phenomena and, if, for example, $\lambda = 0$, the steady state is $x^* = 0.5$. If we suppose a random environment $\lambda_t = \lambda + \sigma\xi_t$ (where ξ_t is a delta correlated white noise, with σ being the intensity of the noise, see Gardiner, 1983) we transform the ordinary differential equation 2.19 into a *stochastic differential equation*, defining a Markovian stochastic process with a stationary probability density. In some cases, such as the one concerning us, the stationary probability density can be found analytically. In our case it is

$$f(x) = Nx^{-1}(1-x)^{-1}\exp\left\{\frac{2}{\sigma^2}\left[-\frac{1}{2x(1-x)} - \lambda\ln\left(\frac{1-x}{x}\right)\right]\right\} \qquad (2.20)$$

where N is a normalization constant. In figure 2.7 we see $f(x)$ for different values of σ.

It is interesting to note that clearly a qualitative change in steady-state behaviour occurs, change unambiguously reflected *in the extrema of the stationary probability density* $f(x)$. This is a very important point, to measure noise-induced transitions we have to measure the extrema of the stationary density (and *not* the mean or other

moments of the probability distribution, see sect. 6.3 of Horsthemke & Lefever, 1984). In the genetic model, for $\lambda = 0$, as we increase the noise intensity σ, $f(x)$ goes from a maximum at $x^* = 0.5$, i.e. the most probable value is the deterministic one, to the appearance of three extrema for $\sigma^2 > 4$, one minimum at $x^* = 0.5$ and two maxima $x_{m\pm}$, so the system now displays bistable behaviour.

2.3.2 Order Parameter for FNNs

In order to get a quantitative theory of a transition phenomenon we have to identify a quantity called *order parameter* that vanishes on one side of the transition and moves away from zero on the other side (Binney et al., 1992). Solé & Miramontes (1995) put forward an order parameter for FNNs: $\Gamma(\rho) = \log_2(N+1) - S(\rho)$ based on the assumption that at low ρ the SK entropy would be maximum due to the practical absence of interactions between the automata. This has been found to be not as general as was supposed, because of at certain values of p_a and in spite of the low ρ, interactions are frequent enough to make $S(\rho)$ have values below its maximum.

As stated above, we would like to study the relation between p_a and the critical density ρ_c. The spontaneous activation represents a sort of multiplicative noise, because of the dependence on the state of the individual (if the automaton is active, there is no spontaneous activation). The "noise level" of the system (number of individuals becoming active spontaneously) depends on the state of the whole system ρ_t^+. Thus, following noise-induced transitions theory, we must find an order parameter looking at the extrema of stationary probability densities. We do not know yet how to find an analytical expression for these densities in the case of FNNs, but we have computed, for various parameters sets, an estimation of the stationary probability density (histogram) for ρ_t^+: $P(\rho^+)$. The stationarity has been assured using the results of section 2.2.3 for the transient length. After the transient has finished, we have used $T = 2 \times 10^4$ time steps to compute the above mentioned probability density. The final result has been obtained averaging over different initial conditions (see figure 2.8)

The computed $P(\rho^+)$ for different values of (p_a, ρ) makes clear the relation between the shape of $P(\rho^+)$ and the dynamical behavior of ρ_t^+. If we have a low density ρ or a low activation probability p_a, ρ_t^+ will be zero, except some irregular bursts of activity that cannot propagate, so the histogram will have a clear maximum at $\rho^+ = 0$. At high enough values of ρ, the rapid propagation of activity resulting from the interaction between automata will make them to remain active almost all the time (if we have not a very low p_a). But at intermediate values of both p_a and ρ, the

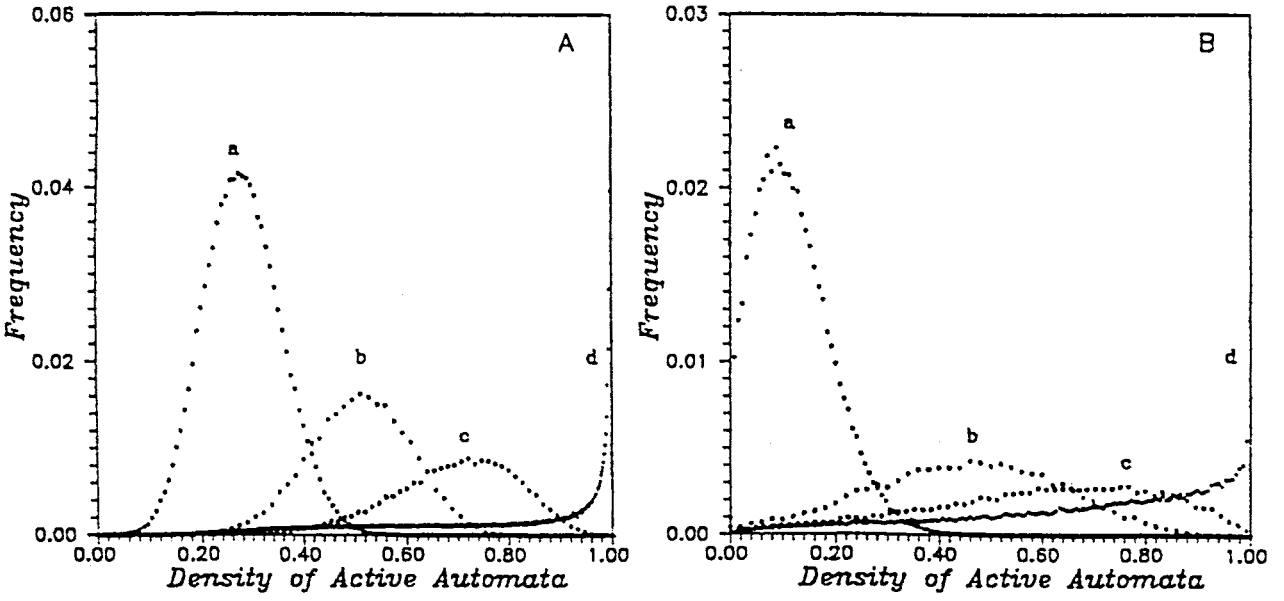


Figure 2.8: Probability densities of ρ^+ with parameters $L = 50$, $S_a = 0.1$, $g = 0.1$, $\theta_{act} = 10^{-16}$. (A) $p_a = 0.01$ and (a)$\rho = 0.05$, (b)$\rho = 0.1$, (c)$\rho = 0.15$, (d)$\rho = 0.25$. (B) $p_a = 0.001$ and (a)$\rho = 0.1$, (b)$\rho = 0.2$, (c)$\rho = 0.25$, (d)$\rho = 0.35$.

interplay between the spontaneous activation and the propagation of the activity will make $\rho_t^+ \in (0,1)$. As we can see in figure 2.8, the qualitative change of shape of $P(\rho^+)$, where the transition occurs, consists of a displacement of the maximum of $P(\rho^+)$ toward the boundary $\rho^+ = 1$, until there is not any extrema in the interval $(0,1)$. In the examples of figure 2.8 we can observe that the above mentioned qualitative change takes place between $\rho = 0.15$ and $\rho = 0.25$ for $p_a = 0.01$ (figure 2.8 (A)(c) and 2.8 (A)(d)) and between $\rho = 0.25$ and $\rho = 0.35$ for $p_a = 0.001$ (figures 2.8 (B)(c) and 2.8 (B)(d)), which was already suggested by the SK entropy in figure 2.5. These phenomena are precisely the type of phenomena for which noise induced transitions theory (Horsthemke & Lefever, 1984) is well suited.

Thus, in FNNs we can define ρ_m^+ such that:

$$P(\rho_m^+) = \max_{\rho^+ \in [0,1]} P(\rho^+) \tag{2.21}$$

Finally we will define the order parameter:

$$\Gamma(\rho, p_a) = 1 - \rho_m^+ \tag{2.22}$$

where obviously the (ρ, p_a) dependence comes from $P(\rho^+)$. In figure 2.9 we can see the order parameter for the FNN whose histograms $P(\rho^+)$ we have seen in figure 2.8. The points where takes place the transition are, for $p_a = 0.01$, i.e. figure 2.9 (A), $\rho_c \simeq 0.215$, and for $p_a = 0.001$, i.e. figure 2.9 (B), $\rho_c \simeq 0.295$, which agrees

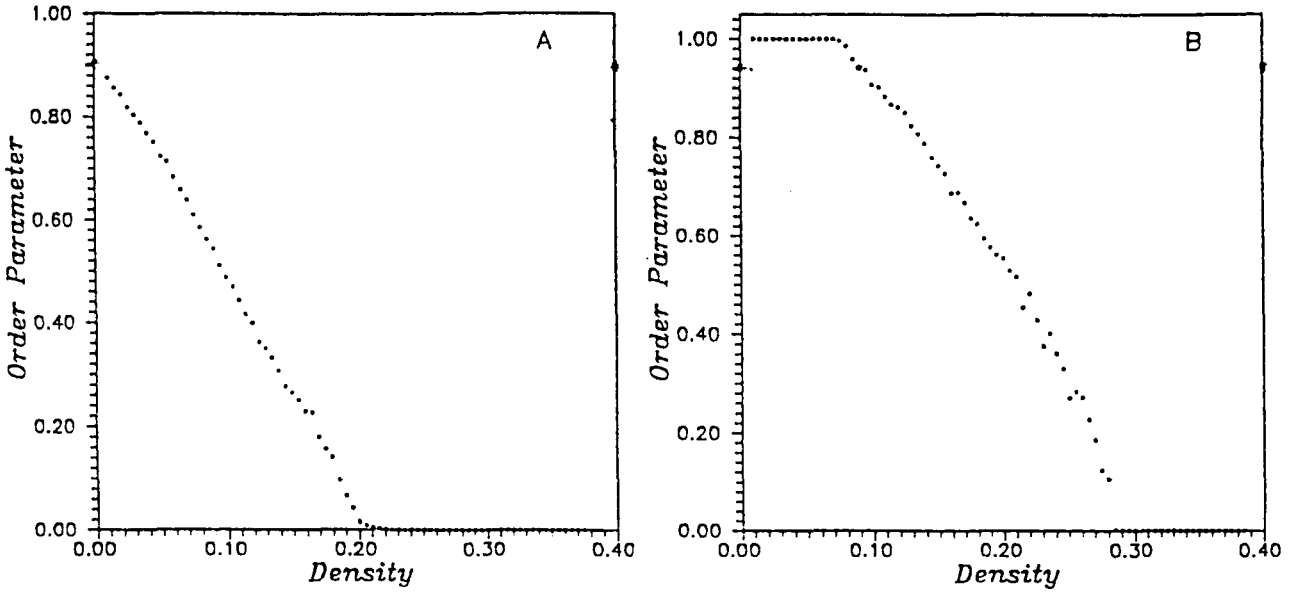


Figure 2.9: Order Parameter $\Gamma(\rho, p_a)$ with parameters $L = 50$, $S_a = 0.1$, $g = 0.1$, $\theta_{act} = 10^{-16}$ and (A) $p_a = 0.01$, (B) $p_a = 0.001$.

accurately with the SK entropy maxima. Once we have an order parameter $\Gamma(\rho, p_a)$ we can explore the parameter space (p_a, ρ) to locate the transition boundaries and to construct a qualitative *phase diagram*, which is shown in figure 2.10. There are three clear areas in the diagram, corresponding to the three shapes that can have $P(\rho^+)$, with the associated three values of $\Gamma(\rho, p_a)$: $\Gamma(\rho, p_a) = 0$ in region II, where we have self-organized oscillatory behavior. This is so because once all the automata are active, they will remain in this state until each automaton becomes inactive again. Since the propagation is very fast, any spontaneous activation will initiate the process again (see figures 2.4(D), 2.4(E), and 2.4(F)). $\Gamma(\rho, p_a) = 1$ in region III where we have almost total inactivity, due to either a very low p_a, where spontaneous activation is a very rare event, or very low ρ, where activity cannot propagate. $\Gamma(\rho, p_a) \in (0, 1)$ in region I. In region I we have an intermediate activity level almost all the time, but with a very irregular behavior (see figures 2.4(A), 2.4(B) and 2.4(C)).

The results obtained as yet pave the way to a hypothesis concerning the functioning of FNNs. On one hand we have the mean time an automaton is active, $\tau'(\rho, p_a)$, that depends essentially on the interactions with other automata, therefore on ρ. On the other hand, we have the velocity of activity propagation, $Y(\rho, p_a)$, that is, the average number of time steps necessaries to reach the state $\rho_t^+ = 1$. These factors, we hypothesize, are enough to find ρ_c: if $\tau'(\rho, p_a) > Y(\rho, p_a)$ the state $\rho_t^+ = 1$ will be reached almost surely, because activity will propagate *before* individuals be-

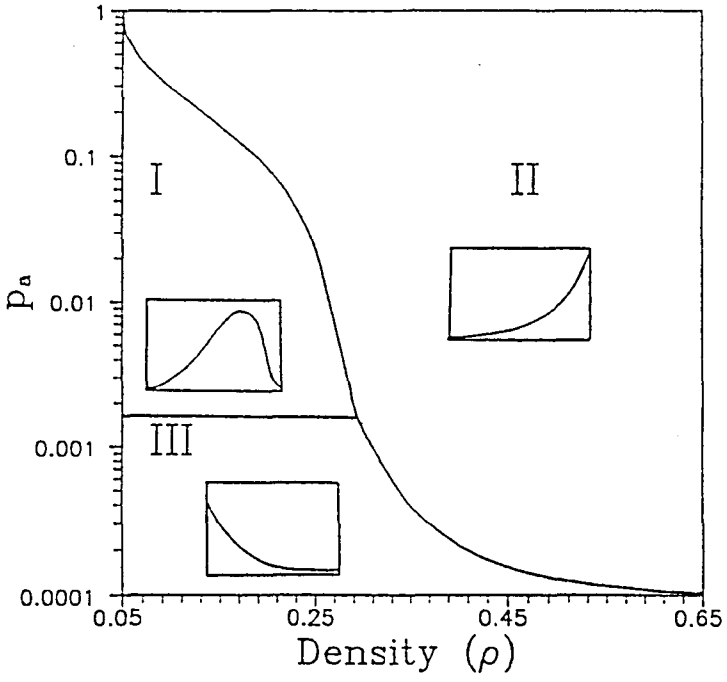


Figure 2.10: Phase diagram computed from $\Gamma(\rho, p_a)$. Parameters are $L = 32$, $S_a = 0.1$, $g = 0.1$, $\theta_{act} = 10^{-16}$. Region I corresponds to $\Gamma(\rho, p_a) \in (0, 1)$. Regions II and III correspond to $\Gamma(\rho, p_a) = 0$ and $\Gamma(\rho, p_a) = 1$ respectively.

come inactive, otherwise individuals will desynchronize due to inactivations before activity propagation is complete. To check this hypothesis it would be necessary to perform some analytic work on FNNs, work that has been carried out simplifying the original FNN definition while retaining the phenomenology.

2.4 Mean field approximation of ρ_c

Some features of the original FNN, as defined in section 2.2.1, could be considerably simplified. We will define the Simple Fluid Neural Network (SFNN) in the following way: We have N individuals $S_i(t) \in \mathbf{R}$ that change their state according to

$$S_i(t+1) = gS_i(t) + g\sum_{j_i^*} J_{ij_i^*} S_{j_i^*}(t) + S_a \Theta[\theta_{act} - S_i(t)] I_i^t \tag{2.23}$$

where $I_i^t \in \{0, 1\}$ with probability $P(I_i^t = 1) = p_a$ and we have made a first order approximation of tanh: $\tanh(x) \simeq x$ removing one of the nonlinearities of the original FNN. The meaning of J_{ij}, S_a, p_a and g is the same as in the original FNN (section 2.2.1). Active states will be defined by $a_i^t = \Theta[S_i(t) - \theta_{act}]$.

What does it mean j_i^*?, it is the *neighborhood.* At a given time step t, the local field $h_i(t) = \sum_{j_i^*} J_{ij_i^*} S_{j_i^*}(t)$ will be computed for all i *before* the change of state

$S_i(t+1)$ is performed. In order to do so, *for each individual* $S_i(t)$, K random connections to some individuals will be established (these individuals will be called *the neighbours*). K is chosen randomly from the distribution

$$P(K=k) = \binom{V}{k} \rho^k (1-\rho)^{V-k}$$

where V is the maximum number of neighbours. This has the same effect as if we threw, at each time step *and* for each element $S_i(t)$, all the N elements upon a $L \times L$ lattice (then $\rho = N/L^2$ will be the density of elements), in order to compute the corresponding local field $h_i(t)$. Thus, we get some kind of "annealed" movement. This is similar to the mean field approximation made in spatially distributed epidemic models (Boccara *et al.*, 1994), where movement was dependent on a parameter m such that the limit $m \to \infty$ was in fact the same as throwing randomly all the elements upon the lattice at each time step. In our case we do so to compute each local field $h_i(t)$, so our system is, in this sense, more disordered.

Considering the density of active individuals at time t

$$\rho_t^+ = \frac{1}{N} \sum_{i=1}^{N} \Theta[S_i(t) - \theta_{act}] \tag{2.24}$$

we can see in figure 2.11 that ρ_t^+ in SFNN has the same temporal behavior as ρ_t^+ in FNN (figure 2.4): Irregular behavior at low densities and more ordered oscillatory behavior for growing ρ. This allows one to apply the FNN order parameter in this case too. We can see $\Gamma(\rho, p_a)$, as defined in section 2.3.2 for FNNs, in figure 2.12, computed for a SFNN.

To sum up, we have a simple FNN, where some nonlinearities have been removed and where each individual, at each time step, establishes connections randomly, as if we had some kind of "annealed" movement. This has simplified considerably the model without loss of interesting behavior because both, FNN and SFNN are phenomenologically identical. In the rest of this section we will explore the relation between the critical density ρ_c and activation probability p_a in SFNN (as we did above numerically with FNN). Throughout this section, the values of the parameters will be $g = 0.1$, $V = 4$, $S_a = 0.1$, $J_{ij} = 1$ for all i, j and $\theta_{act} = 10^{-16}$.

The analysis will be performed for $\rho > \rho_c$, that is, in the region of well developed oscillations. There the behavior of activity spreading is quite well defined: As we see in figure 2.13 the role of spontaneous activation is merely that of starting the process of activity propagation, process that continues by means of interaction among individuals until activity reaches the whole system ($\rho_t^+ = 1$). This would

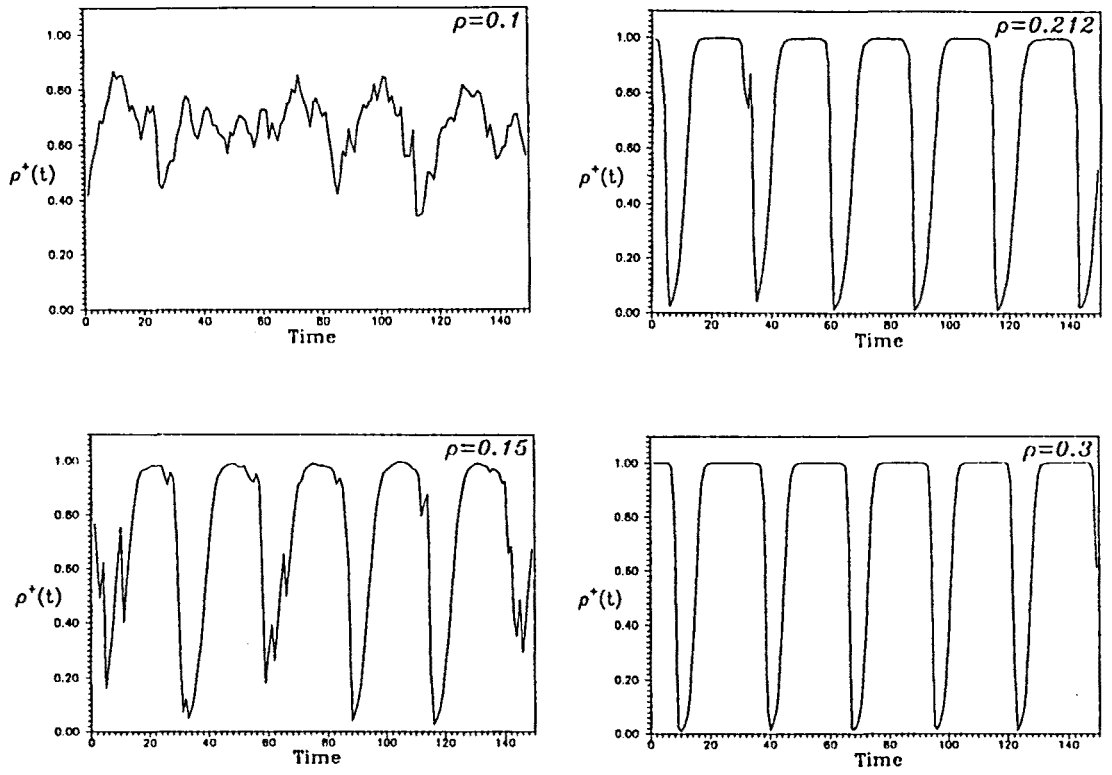


Figure 2.11: Temporal behavior of ρ_i^+ in SFNN, with parameters $L = 100$, $S_a = 0.1$, $g = 0.1$, $\theta_{act} = 10^{-16}$, $p_a = 0.01$ and $V = 4$.

allow us to analyze separately activity propagation and inactivation, assuming in both cases that there is no spontaneous activation.

First of all we will find a condition on V and g to assure the decay of the system to a state where all N elements are inactive. With the above mentioned assumption the evolution for $S_i(t)$ will be

$$S_i(t+1) = gS_i(t) + g\sum_{j_i^*} S_{j_i^*}(t)$$

To see the global evolution of the N individuals we can derive a discrete equation for the state average $< S(t) >= 1/N \sum_{i=1}^{N} S_i(t)$ if we approximate the term $\sum_{j_i^*} S_{j_i^*}(t)$ by the mean field version $V\rho < S(t) >$ so that

$$< S(t) >= (g(1 + V\rho))^t < S(0) >$$

and we can assure activation decay if $g(1 + V\rho) < 1$. If we assume that ρ is as largest as possible ($\rho = 1$), we get the condition

$$g < \frac{1}{1+V} \tag{2.25}$$

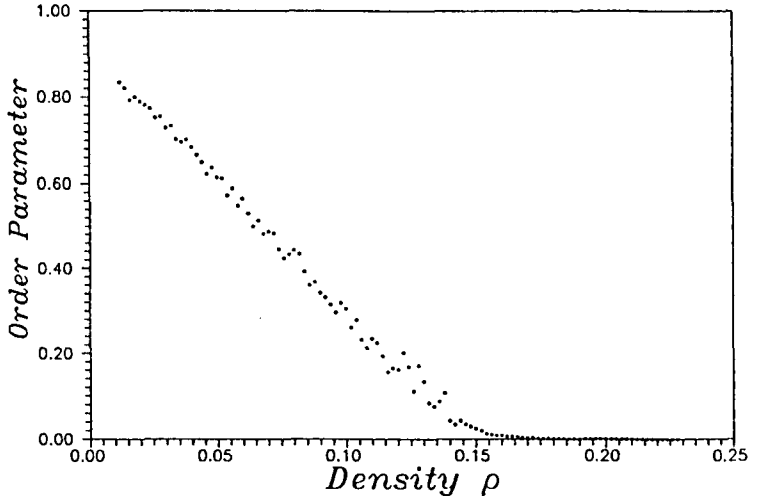


Figure 2.12: Order Parameter $\Gamma(\rho, p_a)$ in SFNN with parameters $L = 100$, $S_a = 0.1$, $g = 0.1$, $\theta_{act} = 10^{-16}$, $p_a = 0.01$ and $V = 4$. Each point has been computed with 10^4 time steps after 2×10^3 transitories.

that is satisfied in our case, because $g = 0.1$ and $V = 4$. Of course, if $< S(t) >$ tends to 0, ρ_t^+ will tend to 0 too. Let us remark that, though $< S(t) >$ tends to zero exponentially but smoothly, ρ_t^+ goes to 0 in very few time steps (as can be seen in figure 2.11 when $\rho > 0.2$).

Now let us study the propagation of activation through the system. In section 2.3.2 we put forward the hypothesis that only two factors were important in order to understand FNN oscillations: the average time $\tau'(\rho, p_a)$ one individual is active between two inactive states and activity propagation $Y_{\lfloor Np_a \rfloor}(\rho, p_a)$, that is, the average number of time steps necessaries to reach the state of $\rho_t^+ = 1$ from an initial state where $\lfloor Np_a \rfloor$ individuals are active, i.e. the mean (integer) number of individuals that would activate spontaneously with probability p_a in a system with all N elements inactive. These are precisely the factors we will analyze in order to compute analitically ρ_c. Intuitively, if $Y_{\lfloor Np_a \rfloor}(\rho, p_a)$ is less than $\tau'(\rho, p_a)$ the state of maximum activation will be reached *before* individuals start the process of inactivation, then we will observe oscillations. So then, ρ_c will be such that

$$Y_{\lfloor Np_a \rfloor}(\rho_c, p_a) = \tau'(\rho_c, p_a) \tag{2.26}$$

Activity spreading can be treated as a branching process if when considering activity by interaction we take into account only the state a_i^t of each individual. In this way, we will say that an inactive individual is activated by its neighbours if there is at least one of them active (it is obvious that this is not the *exact* mechanism by which individuals activate each other, since an individual with all neighbours active, each

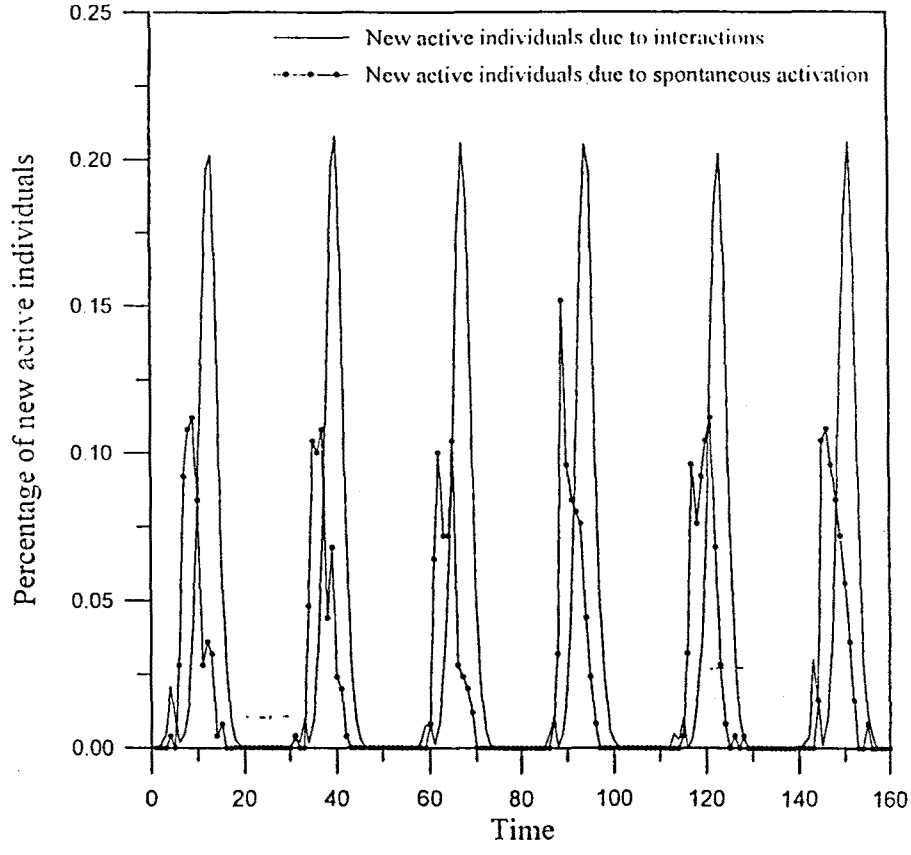


Figure 2.13: Temporal behaviour of the number of new active individuals due to interactions and spontaneous activations. *Values in the spontaneous activation curve are multiplied by* 10. Parameters $L = 100$, $S_a = 0.1$, $g = 0.1$, $\theta_{act} = 10^{-16}$, $p_a = 0.01$, $V = 4$ and $\rho = 0.25$.

one with a very small $S_i(t)$, might not be activated). The probability of having at least an active individual as a neighbour is easy to compute, because of the "annealed movement" we have introduced. If we have i active individuals, the above mentioned probability is

$$\gamma_i = 1 - \left(1 - \frac{i}{L^2}\right)^V \tag{2.27}$$

As we are only considering activity spreading in the oscillations phase, there will be no activity decay, allowing us to compute the probability of j active individuals having i individuals active in the previous time step

$$P_{ij} = P(A_{t+1} = j \mid A_t = i) = \begin{cases} \binom{N-i}{j-i} \gamma_i^{j-i} (1-\gamma_i)^{N-j} & \text{if } i \leq j \\ 0 & \text{otherwise} \end{cases} \tag{2.28}$$

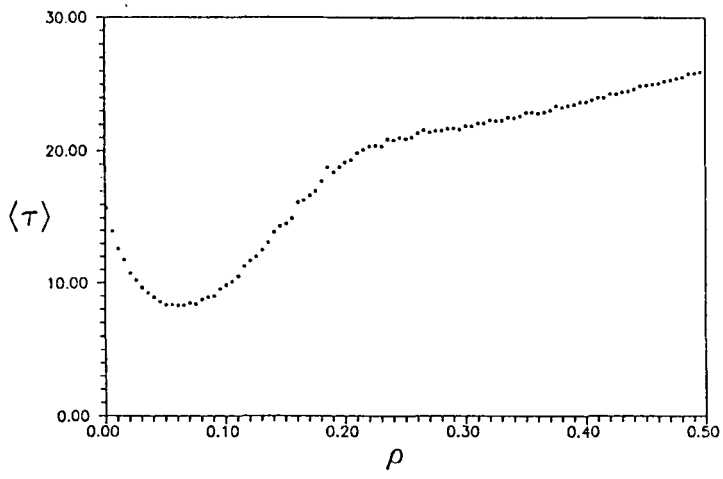


Figure 2.14: The mean time an individual is active between two inactivations. We can see that $< \tau >$ is density dependent with a clear change in the shape of the curve around ρ_c ($\rho_c \simeq 0.212$ in this case). Parameters $L = 100$, $S_a = 0.1$, $g = 0.1$, $\theta_{act} = 10^{-16}$, $p_a = 0.01$ and $V = 4$. $< \tau >$ is computed from 10^3 time steps after 10^3 transitories.

where A_t is the number of active individuals at time t. This defines a branching process that will finish when $A = N$. We will treat this process as a Markov chain (Luenberger, 1979) with stochastic matrix

$$P = \begin{pmatrix} P_{11} & P_{12} & \cdots & P_{1N} \\ 0 & P_{22} & \cdots & P_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & P_{NN} \end{pmatrix} \tag{2.29}$$

with which we can compute the mean number of steps before being absorbed by the unique clossed class of our system, the one element set $\{N\}$. In order to perform the calculations, the P matrix has to be rearranged to get the canonical form

$$P^* = \begin{pmatrix} P_1 & 0 \\ R & Q \end{pmatrix} = \left(\begin{array}{c|ccc} P_{NN} & 0 & \cdots & 0 \\ \hline P_{1N} & P_{11} & \cdots & P_{1(N-1)} \\ P_{2N} & 0 & \cdots & P_{2(N-1)} \\ \vdots & \vdots & \ddots & \vdots \\ P_{(N-1)N} & 0 & \cdots & P_{(N-1)(N-1)} \end{array}\right) \tag{2.30}$$

so that the *fundamental matrix* $M = [I - Q]^{-1}$ of the Markov chain can be found.

The matrix M plays a central role in transient analysis of Markov chains (Luenberger, 1979). M gives immediately the quantity we want to compute. It is easy to

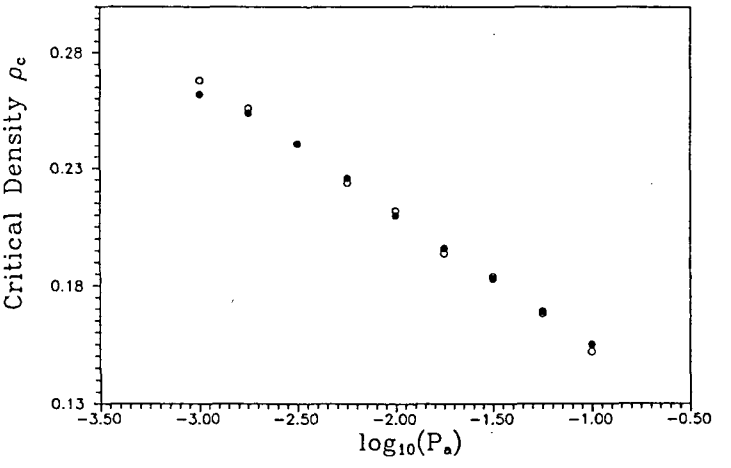


Figure 2.15: ρ_c^{nu} (empty dots) and ρ_c^{an} (filled dots) as a function of $\log_{10}(p_a)$ Parameters $L = 100$, $S_a = 0.1$, $g = 0.1$, $\theta_{act} = 10^{-16}$, and $V = 4$. ρ_c^{nu} is computed from $\Gamma(\rho, p_a)$ with 10^4 time steps after 2×10^3 transitories, averaged over 10 samples.

verify that the ij-th element of Q^k, $q_{ij}^{(k)}$ is the probability of a transition from the state $A = i$ to the state $A = j$ in exactly k steps. The average number of times that starting in state $A = i$ the process reaches state $A = j$ before it leaves transient states and enters the clossed class is

$$q_{ij}^{(0)} + q_{ij}^{(1)} + q_{ij}^{(2)} + \cdots + q_{ij}^{(k)} + \ldots$$

that is precisely M_{ij}, since the identity

$$M = [I - Q]^{-1} = I + Q + Q^2 + \cdots + Q^k + \ldots$$

follows from the fact that Q has all the eigenvalues strictly inside the unit circle (the eigenvalues of Q are $\lambda_j = P_{jj}$ for $1 \leq j \leq N - 1$ and $\lambda_j < 1$ since P is a stochastic matrix). If $\mathbf{1}$ is a column vector whose components are all equal to 1, the mean number of steps before reaching the state of all individuals active, taking as a departure point a state $A = i$, is the i-th component of the vector $M\mathbf{1}$. So, if $Y_j = (M\mathbf{1})_j$, solving the linear system

$$M^{-1}\mathbf{Y} = (I - Q)\mathbf{Y} = \mathbf{1}$$

we can compute any Y_j with the recurrence

$$Y_j = (1 - P_{jj})^{-1} \left(1 + \sum_{i=j+1}^{N-1} P_{ji} Y_i \right) \tag{2.31}$$

Now that we have $Y_{[Np_a]}(\rho, p_a)$ we need to compute $\tau'(\rho, p_a)$. This is a rather difficult calculation and we have approximated $\tau'(\rho, p_a)$ by the inactivation time of a solitary

individual subject to simple perturbations

$$S_{t+1} = gS_t + \xi_t$$

where ξ_t are i.i.d. random variables. These random variables will be such that we should assure that $< S_t > \rightarrow \theta_{act}$, so we will impose $< \xi_t >= (1-\epsilon)\theta_{act}$ where ϵ is a free parameter whose value will be established below. The evolution of $< S_t >$ is easily found (with $S_0 = S_a$) from

$$S_t = g^t S_a + \sum_{j=1}^{t} g^{t-j}\xi_j$$

and it is

$$< S_t >= g^t S_a + (1-\epsilon)\theta_{act} G_t$$

where

$$G_t = \sum_{j=0}^{t-1} g^j$$

Since $g = 10^{-1}$, G_t is easily calculated and it is quite obvious that it can be approximated by G_∞. So then, our estimation of $\tau'(\rho, p_a)$, τ_{per}, will be given by

$$\tau_{per} = (\log_{10} g)^{-1} \left[\log_{10} \theta_{act} - \log_{10} S_a + \log_{10}(1 - G_\infty(1-\epsilon))\right] \qquad (2.32)$$

From equation 2.32 it is clear that ϵ is bounded by

$$\epsilon > 1 - \frac{1}{G_\infty}$$

that is, $\epsilon > 0.1$ and by $\epsilon < 1$, because if $\epsilon = 1$ then $< S_t >= g^t S_a$ and

$$\tau_{per} = \frac{\log_{10}(\theta_{act}/S_a)}{\log_{10}(g)}$$

that is, $\tau_{per} = 15$ for the set of parameters we are working with.

But, how do we determine exactly the value of ϵ? At this point we must resort to the numerically computed $\tau'(\rho, p_a)$ for the SFNN (see figure 2.14). Let $\delta_\epsilon = \epsilon - 0.1$. From equation 2.32 we see that τ_{per} gets larger as δ_ϵ gets smaller, but, looking at figure 2.14, particularly at the region around ρ_c, we see that a τ_{per} much larger than 15 does not make sense. Therefore, we will fix $\delta_\epsilon = 0.01$, a value large enough to keep τ_{per} in the "meaningful" region and small enough to make $\tau_{per} > 15$. Finally, with the set of parameters we have been using and the δ_ϵ above mentioned, $\tau_{per} \simeq 16.95$.

Once we have $Y_{\lfloor Np_a \rfloor}(\rho, p_a)$ and $\tau'(\rho, p_a) \simeq \tau_{per}$ we can find a density ρ_c such that $Y_{\lfloor Np_a \rfloor} \simeq \tau_{per}$. This analytically computed ρ_c^{an} is compared with a numerically determined ρ_c^{nu} using the order parameter $\Gamma(\rho, p_a)$, in figure 2.15. ρ_c^{an} and ρ_c^{nu} have both a linear dependence on $\log_{10}(p_a)$ and agree accurately.

2.5 Self-syncronization and task fulfilment

Once we understand the mechanism of self-generated oscillations, we will change to a more practical point of view and see whether self-synchronized behaviour has any interest to RDAI practicioners. In order to attain that goal, we will introduce a formal framework within which to study the relation between patterns of activity in collectives of simple agents and the ability to fulfil some task. As we have seen in section 2.1, self-synchronized patterns of activity are known to occur in real ant colonies and, though there are some authors that argue against an adaptive interpretation of self-generated oscillations (Cole, 1991a), other authors have pointed out that ordered patterns of activity might be related to task realization by means of either enhancing the possibility of individual information sampling (Robinson, 1992) or interacting with spatial constraints in what concerns the access to the task to be done (Hatcher *et al.*, 1992).

One of the most striking aspects of insect societies is their ability to distribute tasks among individuals in the colony, in such a way that colony needs are completely fulfiled. Of course, no individual can acquire enough information to decide which are the colony shortages, so that task allocation must be a collective property of the colony (see Robinson, 1992 for a review of division of labour in insect societies). How individuals know what task needs to be done at each moment is nowadays a matter of controversy, though there are some hypothesis at hand (Robinson, 1992). One of these, the *fixed threshold model* (FTM), is specially suitable to perform mathematical modelling (Bonabeau *et al.*, 1996) because of its simplicity. Undoubtedly, results concerning task allocation in ant colonies will be of interest to RDAI.

The basic assumptions are that some specific stimulus is associated with each task and that each individual has fixed response thresholds to the various stimuli, so that the lower the threshold the more likely the individual will engage in the task, given exposure. There is an experimental basis that justifies this approach, for example it has been proved the existence of response thresholds in honey bees (see Robinson, 1992, and Bonabeau *et al.*, 1996, for a more detailed discussion on the experimental basis of the FTM)

In this section we will work only with one "abstract", spatially distributed task. Assume that an active (in the sense of "not resting", see section 2.2), though not working, individual perceives, in some way to be specified below, a quantity s of stimulus. It will engage in the task with probability

$$P\left(S_i^{\text{Not working}} \rightarrow S_i^{\text{Working}}\right) = \frac{s^2}{s^2 + \theta^2} \tag{2.33}$$

where θ is the individual threshold associated to the task. Once an individual

is engaged in a determined task, it has some fixed probability p per unit time, independent of the stimulus, of giving up the task. p can be found experimentally (throughout this section, $p = 0.2$, see Bonabeau *et al.*, 1996). We will make some further simplifying assumptions: If an individual is active *and* working, and ceases its work (with probability p) while he is still active, we will consider that a an associated "quantity of task" has been accomplished, identifying this quantity with the stimulus s.

Now we know how individuals activate and engage in work, we have to define a way for distributing a certain, abstract, "amount of task" among the individuals of the system.

Let us assume a $L \times L$ lattice where N individuals are spread on. Each individual will be characterized by a triple $(S_i(t), X_i(t), \theta_i)$ where $S_i(t)$ is the FNN-state of the individual i, $X_i(t)$ is a two-valued variable signaling whether the individual is working ($X_i(t) = 1$) or not ($X_i(t) = 0$) and θ_i is the FTM threshold. Also, a working ($X_i(t) = 1$) and active ($S_i(t) > \theta_{act}$) individual will be doing a certain amount of task $c_i(t)$.

Initially our system will be composed of non-working individuals, with a random initial FNN-state. θ_i will be initially fixed and its value may be either the same for all N individuals or uniformly distributed, between θ_{min} and θ_{max}, among individuals. In this section we will explore both cases. A randomly chosen position of the lattice, say (r_x, r_y) (the "task origin"), will contain a certain amount of total, "abstract ", task $O(0) = C_{in}$ to be performed by individuals. The task origin will be the only square of the lattice having something more than individuals, the rest of squares of the lattice can be occupied by one individual or be empty.

Our system will evolve in time, in what concerns to $S_i(t)$, *exactly* as a FNN (section 2.2), so what remains to be defined is the evolution in time of $X_i(t)$ and eventually $c_i(t)$, that is, the task realization process.

Origin of the stimulus: If an active and non-working individual, say the j-th, is located at the task origin, it will get stimulated by a quantity of task $\alpha O(t)$ with a probability given by equation 2.33. If it becomes a worker, the amount of task it will perform is $c_j(t) = \alpha O(t)$ and a quantity $(1 - \alpha)O(t)$ will remain at the task origin. This rule seeks to stand for a stimulus stemming from a well defined region of space, visited from time to time by individuals of the system.

Effective realization of the task: If an active and working individual i becomes inactive (in a FNN sense) the amount of task $c_i(t)$ will not be taken into account as performed, being added to $O(t)$, the task origin. So, the only way the system has

to effectively perform a certain quantity of task is that an active working individual became a non-worker (with probability p). The quantity of task $c_i(t)$ will disappear from the system and will be considered as done.

Propagation of the stimulus: An active and non-working individual may be stimulated by all its active and working nearest neighbours. Each active and working individual, say the j-th, will be able to provide a quantity of stimulus that will depend on the number of active and non-working neighbours, say n_j. A quantity $s_j^k = \beta c_j(t)/n_j$ will be the stimulus provided by j to each of its n_j non-working neighbours. Thus, an active and non-working individual, say the k-th, will receive a quantity of stimulus

$$S_k = \sum_i s_i^k$$

where i ranges over the active and working neighbours of k. If this individual becomes a worker, with a probability given by equation 2.33 , the quantities s_i^k are substracted from $c_i(t)$ (for all i active and working neighbour of k). If not, nothing happens and the amount of task of the active and working neighbours do not change.

To sum up, a certain quantity of initial task C_{in} needs to be done. This quantity of task will be spread over the individuals, by means of individuals being stimulated by the task origin or by other individuals carrying some amount of stimuli, which eventually would perform the task (it depends on their FNN-state). The unique way to do some part of the task is by changing state from worker to non-worker while being active. In this case the quantity $c_i(t)$ of the corresponding individual will disappear from the system. The quantities C_{in}, α and β are, to some extent, arbitrary, so their value will be fixed at $C_{in} = 50$, $\alpha = 0.15$ and $\beta = 0.1$.

In this section we have explored how certain patterns of temporal activity interact with the process of task fulfilment, depending on the FTM threshold θ_i. In our model, the rules detailed above may be summarized in the following way: "An individual only works if it is active, if it gets inactive before its amount of task is done, that amount of task is not fulfiled". First of all, we have measured how the total task that remains to be done

$$C(t) = O(t) + \sum_{i=1}^{N} c_i(t) X_i(t) \tag{2.34}$$

evolves in time with three different patterns of activity: self-synchronized, non-synchronized with low activity and permanent activity. This has been done for different values of θ, in systems where each individual has the same value $\theta_i = \theta$,

and for a system where each individual has a θ_i chosen at random with uniform distribution in the interval $(1, 10)$.

In general, our results point out that the times taken by the self-synchronized system to get the task done are very near the values obtained on permanent-activity systems. Systems with low activity and desynchronized behave quite worse than the synchronized ones.

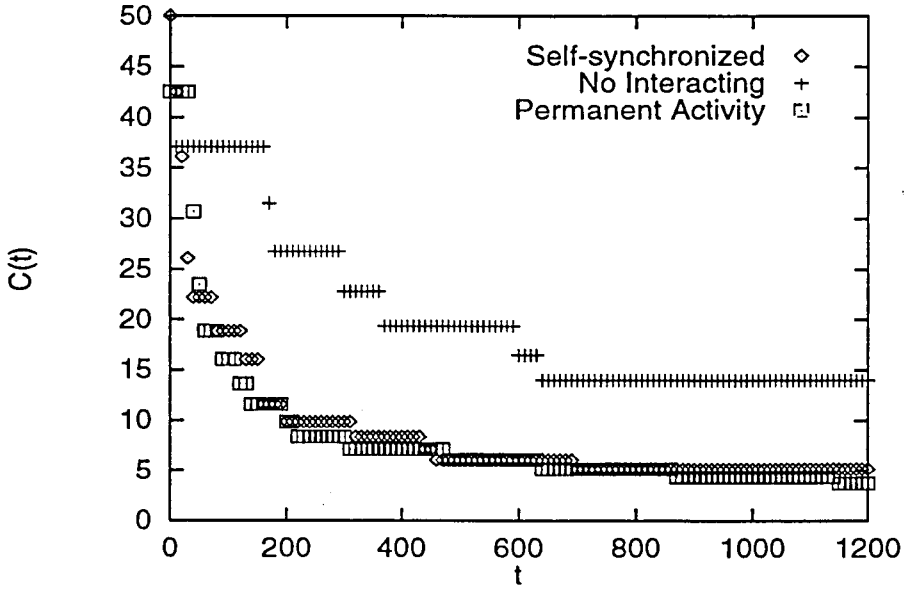


Figure 2.16: Evolution in time of the amount of task that remains to be done at time t: $C(t)$. It is clearly seen that the task C_{in} is accomplished with comparable effectiveness with permanent activity and with self-synchronized activity. See text for parameters, $\theta_i = 5$ for all individuals.

Figure 2.16 shows $C(t)$ for systems with the patterns of temporal activity above mentioned, for a fixed $\theta_i = 5$. It is clearly seen the almost identical temporal evolution of $C(t)$ for self-synchronized and permanent-activity systems, and the worse behaviour of the desynchronized one. A different $\theta_i = \theta$ modifies this behaviour as expected, that is, either increasing or diminishing the time to get the task done, simply because of the different likelihood to get engaged in it. However, with respect to our purpose of comparing the three diferents patterns of activity, the results are qualitatively identical to those of figure 2.16. Figure 2.17 shows the case of different individual threshold, where θ_i has been chosen at random from a uniform distribution between $\theta_{min} = 1$ and $\theta_{max} = 10$. Permanent-activity is slightly better than self-synchronization, being desynchronized systems as bad as they were in the fixed θ case, when compared with the other systems.

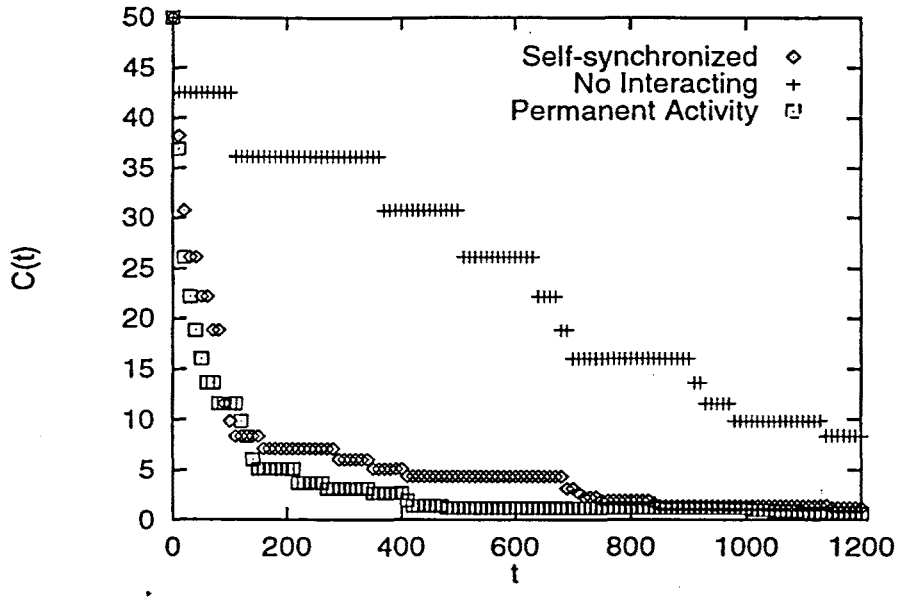


Figure 2.17: In this figure we see $C(t)$ in the same three cases of figure 3, though now θ_i has been drawn randomly from the interval $(1, 10)$

More systematic calculations have been done for systems with fixed θ. We have measured the time a system needs to get the 90% of the task done, that is, the "efficiency" and averaged it over M measures, for $\theta_i = \theta$ ranging from 1 to 7.25. The difference between the efficiency in the three cases we are dealing with increases with θ, as seen in figure 2.18. Self-synchronized activity, for low θ, performs almost as effective as permanent active individuals, being this difference large when comparing desynchronized automata with any of the other behaviors, no matter the value of θ. As θ grows, the difference between self-synchronization and permanent-activity grows, showing that the less likely is to get engaged in the task the more important becomes the mean time one individual is active. Nevertheless, if compared with desynchronized behavior, self-synchronization still has a value near the optimum, that is, near the efficiency of permanent activity.

The phenomenon is not difficult to understand. The key point is the idea that "if it is not active, it does not work" together with the *local* transmission of information (stimulus, task, etc...) from individuals to individuals. If not active permanently, the only way to assure that an individual will have as much *active* neighbours as possible to whom stimulate, and consequently to get the task as scattered as possible is having synchronized activity, assuming, of course, the density is fixed (which is observed in real ant colonies).

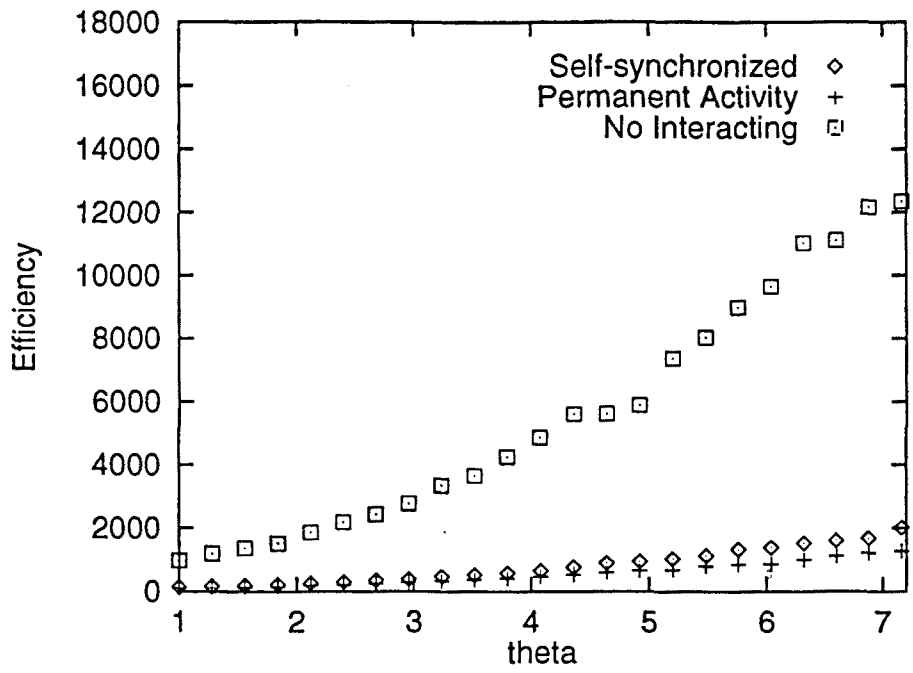


Figure 2.18: Efficiency in accomplishing the task measured for three different temporal behaviours (see text). Efficiency is averaged over $M = 10^2$ measures

2.6 Conclusion

In this chapter we have summarized results on a formal model of self-synchronized behaviour, the *Fluid Neural Network*, by other authors (Solé *et al.* 1993a, 1993b; Solé & Miramontes 1995; Miramontes, 1992; Cole & Cheshire, 1996) and we have introduced some new results (Delgado & Solé, 1997a, 1997c, 1997d), emphasizing the stochastic nature of the FNN fluctuations and the interplay between noise and non-linear interactions. Now, some of the previous questions (section 2.1) can be tentatively answered. Some relevant points are

- The disordered nature of the individual behavior provides a source of (internal) noise which is able, through amplifications by the system, to generate global oscillations.

- The ingredients of ant colony dynamics, as revealed through experiments, together with our model approach, allows us to correctly identify what kind of collective phenomena is at work: Ant colonies operate close to a noise-induced phase transition point. This particular result supports the general conjecture that natural systems use critical points (of some kind) as a way to transfer information optimally (Solé *et al.*, 1996).

- In spite of the earlier hypothesis on ant colony oscillations that considered them as an epiphenomenon (Cole, 1991a), our study clearly points in the direction of a functional meaning. Our results (based on a simple fixed threshold model) indeed show that self-synchronized patterns of activity may behave (almost) as well as a colony with constant activity level. This result provides, for the first time, a consistent understanding of *why* collective oscillations are present.

One of our concrete results is the analytical estimation of ρ_c for SFNNs (up to now the only analytical work carried out on FNN-like systems) though we are currently working on a more accurate calculation of $\tau'(\rho, p_a)$, because the perturbation approach, which needs the introduction of the parameter ϵ whose value must be found heuristically, is not entirely satisfactory. The introduction of interactions with other individuals seems to be necessary for a more sound approximation. Besides, we have introduced an order parameter for the FNN model and we have analysed the behaviour of individuals with spontaneous activations. Our results and some parameters defining the model should be tested with real ant colonies, work which is currently in progress, being performed by B.J. Cole and collaborators (Cole & Cheshire, 1996).

There are other mathematical models of self-synchronization in ant colonies. Different approaches based on ordinary differential equations or probabilistic process algebra have been proposed. These models have been criticized on the basis of a lack of testable predictions or dubious assumptions, see Tofts *et al.* (1992) for a survey. A comparison with FNNs is difficult, due to the different mathematical nature of the models and the different features of real systems emphasized by the models. An example, FNNs puts the emphasis on critical densities and spontaneous activation while the probabilistic process algebra models predict distributions of cycle lengths (Tofts *et al.*, 1992).

At this point, once understood the mechanism underlying oscillations in activity, the "orthodox" way to proceed, from a RDAI point of view, would be to apply self-synchronized processes to the design of algorithms to solve some particular problems (as has been done successfully with stigmergic processes, see section 1.4). This work might begin with the formal framework of section 2.5, solving a "concrete" task instead of an abstract one. However, we have not proceeded this way. Instead, we have turned our attention to more general questions: now that we have a model of social insects behaviour, could this model offer some solution to the *general* problem of collective computation? Is this problem well posed? and, if this is not the case, is there a way to formalize the term "collective computation" in such a way that it might be useful from a RDAI perspective? These questions are the subject of the

following chapters.

Results in this chapter have been published in Delgado & Solé (1997a), Delgado & Solé (1997c) and Delgado & Solé (1997d)

Chapter 3

Computation in collectives of simple agents

3.1 Dynamics and computation

In this chapter we are going to study how Fluid Neural Networks (FNNs) might be able to perform simple computations. As seen in chapter 2, FNNs are spatially distributed dynamical systems, so then a preliminary discussion on how to relate dynamical systems and computation should be appropriate. Most of the discussion will be done in the context of ant colonies, since they are a concrete example of collective system and the source of inspiration of much work in collective computation (Bonabeau & Theraulaz, 1995), including FNNs. We will use, for the moment, the following characterization of *collective* or *emergent* computation (Forrest, 1990): *The constituents of emergent computation are a collection of agents that interact, forming global patterns at the macroscopic level, i.e. epiphenomena, that has a natural interpretation as computation.* Of course, this is not enough to build a theory of emergent computation, but it is to see what kind of phenomena we are dealing with in the case of collectives of simple agents. From real experiments with ant colonies, it is easy to ascertain which properties are characteristic of computation by collectives of simple agents (Bonabeau & Theraulaz, 1995; chapters 5, 6 & 7): it is distributed, i.e. there is not a central control governing the actions of individuals; it is emergent, as was discussed above and it is noise induced (Deneubourg *et al.*, 1986). Perhaps the last point deserves further explanation. There are some computational phenomena which could be called "noise induced", for example associative memories that need noise to work properly (noise in Hopfield neural networks makes unstable some spurious patterns, favouring the recovery of stored patterns (Amit, 1989)), but this "simulated annealing" sense of noisy computation is not what we are talking about. What we mean by noise induced computation is the enhancement of computation

due to the noisy nature of the system, simply put, without noise there is no computation. As far as we know, there is no computational theory to deal with this kind of phenomenon, although the problem of computation *despite* noise (unreliability) has attracted the attention of computer scientists since von Neumann's work (von Neumann, 1956; Gacs & Reif, 1988).

3.1.1 Dynamics, but...

Now we are going to review three collective behaviours characteristic of almost all ant species. These behaviours, collective decision making, collective sorting and task allocation, are representative of the kind of phenomena known as emergent computation, according to the characterization mentioned above. The *natural* interpretation as computation will be obviously the role played by these behaviours in colony survival, that is, the functional properties with which the colony is endowed because of these behaviours. This small review will be useful also in the next chapter, when we discuss alternatives to the relations between dynamics and computation seen in this chapter.

Collective decision making

Experiments with some species of ants (Beckers *et al.*, 1990) have pointed out their collective ability to choose the richest out of several food sources of different quality, showing capabilities far beyond those of an isolated ant. To be more specific, two food sources were presented to an ant colony, simultaneously and at the same distance of the nest, to measure the patterns of source exploitation. With two equal sources, the response of the colony was to choose one of them, usually the one that was first discovered. If the sources were different and were discovered simultaneously, the richer one was the more exploited. However, if a source was first discovered, the introduction of a richer source once the trail to the first source is well formed induces distinct behavior depending on ant species. Trail recruiting ants (for example, *Lasius niger*, *Iridomyrmex humilis* or *Pheidole pallidula* (Hölldobler & Wilson, 1990)) are unable to switch in order to exploit the richer source, while group/trail recruiting ants (for example *Tetramorium caespitum* (Hölldobler & Wilson, 1990)) exploit instead the richer source no matter the order of discovery. To sum up, if two different sources are discovered simultaneously, the richer one becomes always much more exploited, but a late introduction of the richer source may result in switching depending on recruiting systems, that is, on ant species.

In the same paper where the experiments are described (Beckers *et al.*, 1990), the authors propose a mathematical model: Let us assume we have N foragers of

which X_i are at the source i ($i = 1, 2$), E are lost recruits and $N - E - X_1 - X_2$ are in the nest to be recruited. Recruitement accuracy is formalised by means of the fraction of ants in the nest that will become foragers f_i. The model is

$$\partial_t X_i = a_i X_i f_i (N - \sum_{i=1}^{2} X_i - E) - b X_i + cE \quad (i = 1, 2)$$

$$\partial_t E = (N - \sum_{i=1}^{2} X_i - E) \sum_{i=1}^{2} (a_i X_i (1 - f_i)) - pE - 2cE$$

where constants have the following meaning: Lost foragers can return to the nest every $1/p$ time units on average or can find one of the two sources every $1/c$ time units on average. $1/b$ is the average number of time units that ants stay at the source. The fraction f_i allows one to take into consideration the differences between recruitment modes. Trail recruitment depends on the number of foragers at the source, so it is reasonable to make it depend on X_i. Moreover, because the more individuals are at the source the more reinforced is the trail, a monotonic increasing function $g(X_i)$ would be a right choice for f_i. Group/trail recruitment have the same features as trail recruitment but a constant term has to be added, because of the recruitment due to direct contact among individuals $g(X_i) + k_i$.

The model is able to reproduce accurately the experimental results. The cause of the symmetry breaking phenomenon when we have two identical sources is clearly seen in the model: trail recruitment and group/trail recruitment have both an autocatalytic term amplifying small differences in initial conditions (which source is discovered first). However, the constant term in group/trail recruitment makes a difference when the two sources are different and not simultaneously presented (the richer the last). In this case the model with group/trail recruitment has just one stationary state, i.e. exploiting the richer source, whereas trail recruitment has several stationary states and which one will be reached depends on initial conditions.

In a remarkable study, Bonabeau has shown that a simple model of cooperative food retrieval in ant swarms can exhibit flexible behavior when close to the instability point described by the previous type of models. Close to instability structured patterns of activity can grow and be maintained (Bonabeau, 1996; see also Bonabeau, 1997). In a related work, another model of this phenomenon was introduced by Mark M. Millonas (Millonas, 1992). Its departure point was the movement of ants along paths where some quantity of scent was laid, following an experimentally determined probabilistic law to switch at bifurcations (Deneubourg *et al.*, 1990). This allows one to define a microscopic dynamics of ants moving on an ant network, that, under some restrictions of homogeneity and locality, leads to a set of kinetic equations. Let

V_i be the length of the i segment. An ant at a vertex of the network, coming from segment i, will choose segment f to continue its walk with probability $W_i^f \propto w(\sigma^f)$ depending only on the quantity of scent σ^f present at segment f. This transition probability will be of the form $w(\sigma) = (\alpha + \sigma)^\beta$ where the β paramater plays the role of an inverse noise parameter and α is a bias in the scent. These parameters are species dependent (Deneoubourg *et al.*, 1990). For technical reasons it will be written $w(\sigma) = e^{-\beta\epsilon(\sigma)}$ where $\epsilon(\sigma) = -\ln(\alpha + \sigma)$. Defining W_i^f in this way, a very important condition is satisfied, that of *detailed balance*: $W_i^f e^{-\beta\epsilon_i} = W_f^i e^{-\beta\epsilon_f}$. The next step is to define a master equation (Gardiner, 1983) for the density of ants at segment V_i

$$V_i \partial_t s^i = \sum_j w(\sigma^i)s^j - w(\sigma^j)s^i \tag{3.1}$$

On the other hand ants lay scent on the segments, which can be written in the following way:

$$\partial_t \sigma^i = -\kappa\sigma^i + \eta_i s^i \tag{3.2}$$

where σ^i is the scent density in segment i, κ is the evaporation rate and η_i is the scent quantity laid by an ant per time step in segment i. Equations 3.1 and 3.2 define the evolution of ants and scents on the network. However, to be able to solve completely these equations, one more assumption has to be made: Separation of time scales. We will assume that temporal evolution of scent occurs in a time scale much slower than that of ant's density, so that constant scent in ants temporal evolution can be supposed. This is called *slaving* after H. Haken (see chapter 1, Haken, 1977). In such a case, ants will relax to their equilibrium values before scents change much and s^i can be replaced in equation 3.2 by their equilibrium value s_e^i which can be computed thanks to the property of detailed balance (Gardiner, 1983)

$$s_e^i = \frac{N(\alpha + \sigma^i)^\beta}{Z} \tag{3.3}$$

where N is the number of ants and

$$Z = \sum_i V_i(\alpha + \sigma^i)^\beta$$

is a normalization factor. Inserting 3.3 in 3.2 a nonlinear evolution equation for scents can be obtained

$$\partial_t \sigma^i = -\kappa\sigma^i + \eta_i \frac{N(\alpha + \sigma^i)^\beta}{Z}$$

The case concerning us, that of collective decision, can be rephrased in terms of this model considering two equal segments whose point of departure is the nest, having in their ends food sources. The parameters defining the problem will be η_1 and η_2. This means that ants returning from each food source lay scent at rates according to the quality of the food source (in agreement with experiments). Using the parameter $R = (\alpha + \sigma^1)/(\alpha + \sigma^2)$ and some algebraic arrangements, finally the following equations are obtained

$$\partial_t \sigma^1 = -\kappa\sigma^1 + \frac{\eta_1 N}{V(1 + R^{-\beta})}$$

$$\partial_t \sigma^2 = -\kappa\sigma^2 + \frac{\eta_2 N}{V(1 + R^{\beta})}$$

The analysis of this model puts the emphasis in the number of ants N instead of recruitment strategies, as did the previous model. Defining the parameter $\gamma_i = N\frac{\eta_i}{\alpha\kappa V} \propto N$ (see (Millonas, 1992) for details) it is found that, depending on N, different behavior concerning the switching ability is obtained. First, let us consider $\eta_1 = \eta_2$. In this case, for $\gamma < \gamma_c$ both sources will be exploited evenly, but a supercritical bifurcation (Wiggins, 1990) occurs at γ_c, so that for $\gamma > \gamma_c$ one of the food sources is chosen, depending on random fluctuations in the initial density of ants s^i. On the other hand, for $\eta_1 \neq \eta_2$, exploitation of the better food source is a stable solution for all values of γ, although above a certain γ_c there will be another stable solution, i.e. to exploit the poorer food source. This will define two disconnected branches of stability. However, if the number N of ants grows from zero, as usually happens in experimental situations, the better food source will be always the more exploited. Formally, this is due to a transcritical bifurcation (Wiggins, 1990).

Collective sorting

In Deneubourg *et al.* 1991 some experiments are described concerning sorting and clustering in ant colonies. In one of them a large number of ant corpses were placed randomly on a arena. A rapid response of the colony (*Pheidole pallidula*) was observed, sorting the corpses and clustering them. At the end of the process one or two clusters of corpses remained on the arena. The other experiment consisted of larvae sorting. The larvae of a colony of *Leptothorax unifasciatus* were tipped out onto an area. Workers brought back larvae to the nest making some small piles. After a while, one large pile of larvae was observed, with a clear separation between small larvae and large larvae.

In the same paper a model of these phenomena was introduced. It is not a model made out of differential equations, as the models we have seen above, but a discrete model in time and space, a kind of mobile cellular automata. We assume an area of $L_1 \times L_2$ squares, each one containing an individual and/or an object that can be of one of two types, A and B. Individuals are able to move randomly in one of the four directions of the Von Neumann neighborhood, except for the borders of the area. It has, therefore, zero-flux boundary conditions. Each individual can hold an object if he is located at the same place and if he is not holding another object. An individual is also able to put down an object, if he is holding one. An object will be picked up depending on the spatial environment of the object. Each individual has a short memory where he can remember how many objects of type A he has found in the last m steps, and also how many objects of type B. Thus, an individual is able to compute an estimation of the ratio of A objects, f_A^m, and of the ratio of B objects, f_B^m, in the environment of a given object, that located at the same coordinates of the individual at that time step. To pick up an object is done in a probabilistic manner

$$P_{A/B}^{up} = \frac{(K^+)^2}{(K^+ + f_{A/B}^m)^2}$$

where K^+ is a constant. To put down an object we define

$$P_{A/B}^{down} = \frac{(f_{A/B}^m)^2}{(K^- + f_{A/B}^m)^2}$$

where K^- is a constant. As is obvious from the definition of $P_{A/B}^{up}$, the more objects an individual finds of a certain type the less probable is that he picks up one of them, and from the definition of $P_{A/B}^{down}$ we see that an individual will put down an object with a probability that increases with the objects of the same type he finds while walking around. This simple model reproduces the results observed in experimental manipulations, at least qualitatively (Deneubourg *et al.*, 1991). In figure 3.1 we can see snapshots of the object locations at different time steps. An increasing clustering according to the type of the objects is observed. Remarkably, this model attains a global result (clustering of objects initially spread out on the lattice) using only local rules, just as did Millonas' model of collective decision.

There are some refinements of this model, introduced in (Gutowitz, 1993). The individuals, as defined up to now, are called "basic" ants; he defines alternative rules for movement, picking up and putting down objects. These ants are called "complexity-seeking" ants. Roughly, these ants are able to measure what Gutowitz calls the local complexity of a square of the lattice. This is a measure of how many consecutive changes of content there are among neighbor cells, i.e., the number of

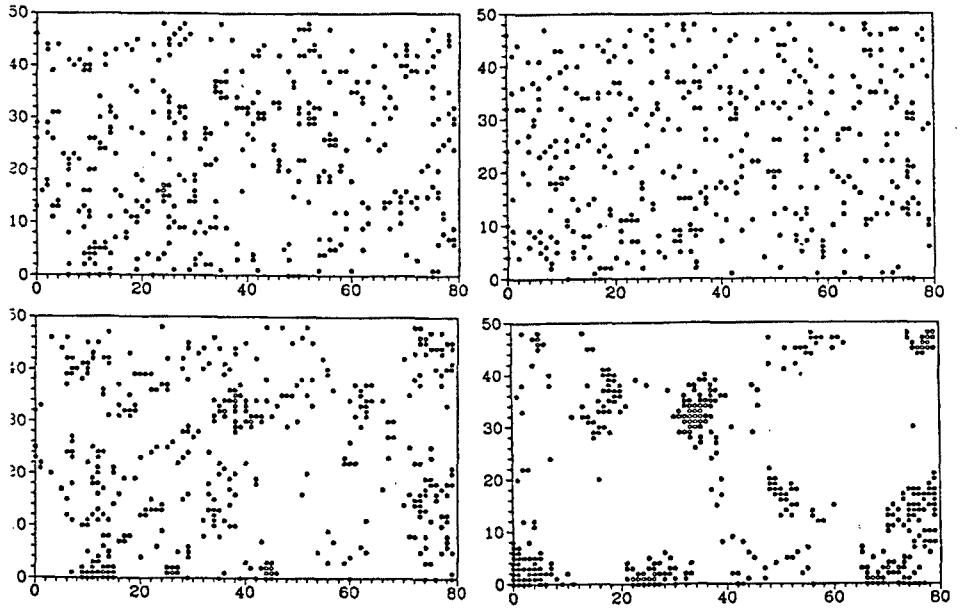


Figure 3.1: Four snapshots of the clustering process in "ant-like" robots (not shown in the figure). Each snapshot shows the object distribution in a 50×80 lattice. The order in time is from top to bottom and from left to right.

"faces" that separate cells of different type. This determines the direction of the movement and manipulation of objects. Complexity-seeking ants are able to achieve the same resuts as basic ants, but faster. This is seen in the decreasing rate of the spatial entropy (see Gutowitz, 1993, for details).

Task allocation

Division of labor is one of the most explored aspects of behavior in ant colonies. Every instant, active individuals in ant colonies are performing one of the essential tasks for colony upkeep: Foraging, patrolling, nest maintenance or midden work. Every task is done accurately by a certain number of specialized individuals. This specialization is unquestionable in what concerns some tasks such as reproduction, where there are clear differences, morphologically and in reproductive status, among queens, workers and males (Hölldobler & Wilson, 1990). The term "caste" was used originally to distinguish among these classes of individuals. However, later on "caste" has been used to distinguish among workers performing different tasks. In polymorphic ant species, where ants morphologically different carry out different tasks, we have "physical" castes (*physical polyethism*) and in monomorphic ant species we have "temporal" castes, where task is age-dependent (*temporal polyethism*). Underlying this notion of caste there is the idea of a *fixed* specialization of individuals, which would determine colony behavior through *caste*

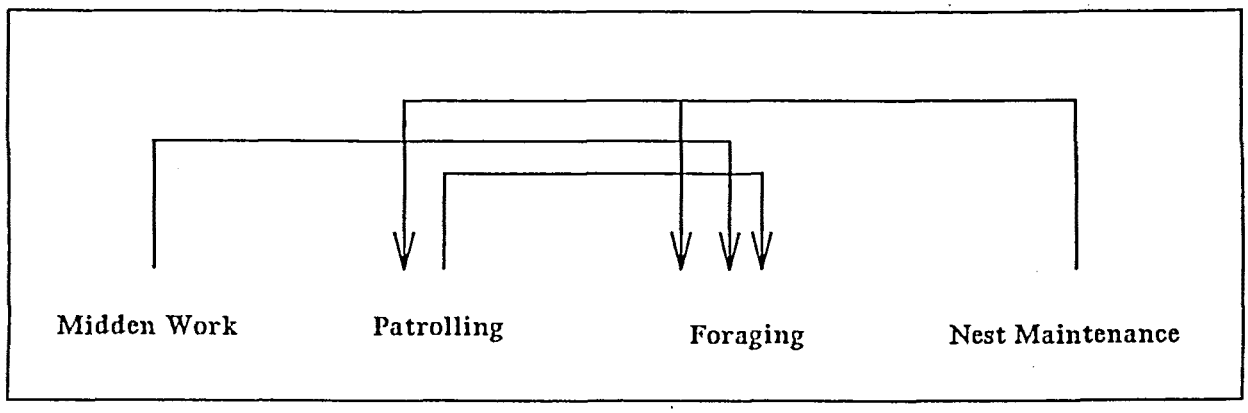


Figure 3.2: Perturbations in *Pogonomyrmex barbatus* colonies induce activity switching (see text). Arrows in the figure show for each activity, which other activities regular workers will do (after Gordon, 1989b)

distribution (Oster & Wilson, 1978) caste distribution being a species character subject to evolutionary processes.

This fixed specialization of workers is nowadays controversial, because, as argued by some authors (Gordon, 1989a), there is a large body of evidence that individuals switch tasks, in both polymorphic and monomorphic ant species. It is interesting that, as Gordon points out, these notions of fixed specialization and caste distributions could have been reinforced by experiments done with small number of colonies in very stable conditions and over short time periods. Concerning task allocation and division of labor, we will assume the ideas unfolded in (Gordon, 1989a), where a complex dynamical process between the individual level and the collective level is the ultimate reason for the percentage of individuals performing each task. For other current points of view concerning division of labor, see (Robinson, 1992). In which sense can we say that task allocation is a computational process? It is obvious if we pay special attention to one of the more interesting properties of task allocation in ant colonies: their *flexibility*. Usually the environment of an ant colony is, to some extent, unpredictable and the colony has to deal with very diverse situations throughout its life. These situations include perturbations in the number of individuals doing some task and task allocation systems should compensate this loss, so that colony upkeep continues as unperturbed as possible. This compensation can take the form of either reassignment of individuals to tasks or activation of inactive individuals or both, involving, among other things, task switching. But this is not so simple. Some experiments (Gordon, 1989b) show that task switching alone does not account for some colony-level phenomena. The experiments with colonies of Harvester ants (*Pogonomyrmex barbatus*) consisted of marking exterior workers (those either foraging, patrolling, doing nest maintenance or midden work)

to compare their behavior in unperturbed colonies with their behavior in colonies where some task has been intensified. Furthermore, colonies of different age were also compared with respect to task fidelity. Results seemingly suggested some level of task fidelity, although perturbations made workers change task. This change is not independent of the task an individual is performing, since some characteristic trends, such as the tendency of workers to switch task out of nest maintenance and into foraging, were found; see figure 3.2 to sum up the changes among tasks in unperturbed workers. It was also observed that task fidelity is weaker in younger colonies. There are also very interesting results concerning colony-level dynamics suggesting that there is not a simple relation between task switching and number of workers engaged in different tasks. For example, in the perturbation experiments foragers rarely switched tasks to do nest maintenance but perturbations increasing nest maintenance work decreased the number of workers foraging.

From these and other experimental results (Gordon, 1989a; Gordon 1989b; Hölldobler & Wilson, 1990; Robinson 1992 and Wilson, 1971) it is clear that the interplay between the colony and the individual, in what concerns task allocation, cannot be accounted for either by optimization models (Oster & Wilson, 1978) (since we have dynamical task switching) or by simple linear relationships between numbers of workers engaged in such and such task. Recently some models have been proposed to explain these phenomena, having in common a serious consideration of nonlinearities as a possible cause of the complex dynamics underlying these task allocation systems. A model of the above mentioned experiments was suggested in (Gordon *et al.*, 1992) based on attractor neural networks (Amit, 1989). Individuals are divided into eight categories according to task and activity: Forager, patroller, nest maintenance worker and midden worker, being either active or inactive for each task. A triple $(\pm 1, \pm 1, \pm 1)$ represents each category. Interaction among ants is defined through three matrices α, β and γ, each one representing one binary decision, so that one of the eight categories can be chosen from the interaction with other ants. These matrices are defined according to some plausible rules of interaction among workers. If we have an ant in a state (a_k^t, b_k^t, c_k^t), the next state will be given by

$$a_k^{t+1} = \Theta(\sum_{j \neq k} \alpha_{kj} a_j^t)$$

$$b_k^{t+1} = \Theta(\sum_{j \neq k} \beta_{kj} b_j^t)$$

$$c_k^{t+1} = \Theta(\sum_{j \neq k} \gamma_{kj} c_j^t)$$

where $\Theta(x) = 1$ if $x > 0$ and $\Theta(x) = 0$ otherwise. The existence of a global attractor for this system can be demonstrated (see (Gordon *et al.*, 1992) for details) using the energy function

$$E = -\frac{1}{2}\left[\sum \alpha_{jk} a_j a_k + \sum \beta_{jk} b_j b_k + \sum \gamma_{jk} c_j c_k\right]$$

Multiple attractors can be obtained with the same system modifying the γ matrix (this can be shown with the same energy function). Although the model does not reproduce exactly either the interactions among workers or the structure of these interactions, the type of dynamics proposed reproduces, at least qualitatively, the phenomena to be explained, that is, the non-linear behavior of workers engaged in some tasks with respect to perturbations in workers performing other tasks.

There are other factors, besides that of interaction among workers, affecting allocation of individuals into different tasks. One of these factors is colony size (Hölldobler & Wilson, 1990). Recently a model based on ordinary differential equations and stochastic processes has been proposed to account for this phenomenon (Paccala *et al.* 1996). Briefly, the model is able to suggest how individuals have to regulate the per capita rate of social interaction in order to keep in balance the acquisition of information through environmental stimuli and social exchanges. Besides, the comparison between deterministic and stochastic versions of the model is useful to see the plausibility of the deterministic model for small groups, perhaps surprisingly due to the collective nature of the phenomenon. Finally, the model is able to show how a very simple mechanistic interaction allows the system to achieve distributions near those that would maximize colony fitness.

3.1.2 ...where is the computation?

Usually, in order to understand computational phenomena in collectives of simple agents researchers have put the emphasis on modelling the dynamics of these phenomena, showing how to apply that dynamics to the solution of some problem, typically optimization problems such as the Travelling Salesman Problem (Bonabeau & Theraulaz, 1995, chapters 6 & 7; Colorni *et al.*, 1992; Dorigo *et al.*, 1996). But if we want (and we do) a theory of emergent computation we *must* relate computational processes in collectives of simple agents to well known classical notions of computation (Hopcroft & Ullman, 1979). We have been reviewing some experimental and theoretical results concerning three different behaviors of ant colonies that may be called "computational", that is, emergent behavior with some well defined utility for the colony. However, what we have seen are models of the *dynamics* of the system, not of the *computational* properties. From those models it is relatively

easy to construct algorithms to solve problems like ants do, that is, in parallel and using nonlinear and/or noisy distributed mechanisms. But, what if we want to know about these phenomena *computationally*?.

One answer could come from the use of theoretical computer science tools to model the dynamics of the phenomena we want to study. Thus we would have from the very beginning a model in computational terms, with which to obtain information concerning dynamics *and* computation. This has been done using the formalism of *Probabilistic Process Algebra (PPA)* (Tofts *et al.*, 1992; Tofts, 1993; Tofts, 1996; for a definition of probablistic process algebra, see Tofts, 1994), a probabilistic extension of process algebra (Milner, 1990). PPAs are a formalism to build up general concurrent systems made of interacting individuals, hence their usefulness in modeling collective behavior since that is, generally speaking, what collectives of agents are. We will not detail either process algebra or PPAs, but we will try to offer some idea of the formalism, summing up the experiments done by A. Sendova-Franks and N.R. Franks on colonies of *Leptothorax unifasciatus* ants (Sendova-Franks & Franks, 1993) and the theoretical PPA models of C. Tofts (Tofts, 1993). The experiments, very briefly, give as a result a correlation between age and task (temporal polyethism), though this correlation was found to be extremely weak. To explain this and other results, such as the rough correspondence between spatial location and task, Sendova-Franks & Franks (1993) proposed the "foraging-for-work" hypothesis, i.e. ants actively seek tasks to perform. Under this hypothesis, correlation between age and task would be a by-product (which is best known as a side effect, or emergent property, considering the system we are dealing with here (Forrest, 1990)) of the task allocation system. To check this hypothesis, Tofts (1993) introduces an algorithm for task allocation based on "foraging-for-work". This algorithm is formally defined in terms of a PPA called "Weighted Calculus of Communicating Systems" (WSCCS; Tofts, 1994) which allows each agent to be defined as a probabilistic process and compose them in parallel. The algorithm, roughly, operates according to the following idea: If an individual could not successfully perform its current work, it seeks to work in one of the tasks nearer to the one it is performing. It is a computational model with which computational properties can be measured, such as the number of states of the system, and dynamical properties demonstrated, such as the stability of states consisting of an equal number of individuals distributed among all tasks. Finally, the appearance of an age structure is observed in the model, as observed experimentally (unfortunately there is not enough space to consider in detail the model, see Tofts, 1993). The use of PPAs to model biological phenomena is very promising, though some points remain to be worked out, i.e. the importance of

space in biological phenomena is well-known (see, for example, Deneubourg, 1977) this being a problem not definitely solved in process algebra (Baeten & Bergstra, 1991). Furthermore, there are some practical difficulties arising when dealing with systems with a large number of elements (Tofts, 1996), a property that must be taken into account in collective systems.

Another possibility would be to consider dynamical systems modeling collective behavior *and* computing[1], according to some definition of computation in dynamical systems; for example, a dynamical system computes if it can simulate some formal computing device, such as deterministic automata or Turing machines (Hopcroft & Ullman, 1979): There are systems of ordinary differential equations simulating universal Turing machines (Branicky, 1995), though these systems are quite ad-hoc and it is difficult to use them to model the behavior of real systems. Other systems capable of simulation of Turing machines are cellular automata and neural networks (Garzon, 1995) but the above criticism also applies in this case. A feasible alternative would be to compute by means of embedding universal logical gates (logical gates, such as the NOR gate, with which to build up any other logical gate) in dynamical systems: Universality in the game of "Life" has been proved in this way (Berlekamp *et al.*, 1985). This will be the approach followed here to explore the computational capabilities of a slightly modified version of FNNs. However, a relevant difference must be pointed out in relation with deterministic discrete (-state) dynamical systems like the Game of Life (GL). Although a NOR gate can be built up by means of an appropriate set of initial conditions plus the GL rules, this gates -as in other deterministic systems- are highly sensitive to perturbations. More precisely, any small amount of noise able to modify a single bit of the NOR gate will destroy it. This fact is not taken into account in the standard theory. Following our previous approach based on stochastic dynamics, we could ask ourselves if a feasible NOR gate could be properly implemented in such a way that noise was not relevant. Still more important, given the relevance of phase transtitions in collective dynamics, perhaps a NOR gate could be obtained if an appropriate FNN is constructed where probabilistic individuals could take advantage of phase transition points as fluctuation amplifiers in order to process information.

[1]Note that we are now making a distinction between solving a *particular* problem and *general* computation, in the sense of the abstract theory of computation

3.2 Universal Computation in FNNs

The modified FNNs are defined as follows: We consider a set of N entities (which we will call "automata" or "ants") defined, in the general case, by a set of elements each one represented by m bits, i.e. $\mathbf{S}_i(t) \equiv (S_i^1(t), ..., S_i^m(t))$. Here we take $S_i^j \in \Sigma \equiv \{+1, -1\}$. For simplicity, we consider single-bit individuals (i. e. $m = 1$). The density of automata will be indicated by $\rho = N/L^2$. Each time step, all elements will move at random towards one of the eight nearest lattice positions (if available).

The dynamics of this network is described by means of a transition rule defined by the transition probability:

$$P(S_i \to -S_i) = \frac{1}{2}\left[1 - \tanh\left(\beta \left(h_i(t)\, S_i(t) - \Theta_i\right)\right)\right]$$

where h_i is the local field, defined as:

$$h_i(t) = \sum_{j \in B_i(\rho)} J_{ij} S_j(t)$$

being the sum extended over the "neighbors" of $S_i(t)$. Here the strength and type of connection is the same as in section 2.2.1, with the obvious changes to deal with the different representation of states.

For the particular case where $\lambda_{ij} = 1$ for all cases (ferromagnetic FNN) it is not difficult to show that the macroscopic state defined by $m(\{\mathbf{S}\}) = \sum_j S_j/N$ evolves following the mean-field equation:

$$\Gamma \frac{\partial m}{\partial t} = -m + \tanh\left[\beta\rho(m + h)\right]$$

where h is the so called external field and Γ is a constant rate, defining the characteristic time scale of relaxation. We have not provided details on how we derived this equation since the techniques used have been the same used to derive equation 1.6 (chapter 1). For $h = 0$, the stable attractors of this system will be: (i) $m_0^* = 0$ for $J \equiv \beta\rho < 0 = J_c$ and $m_\pm^* \neq 0$ (with $m_+^* = -m_-^*$) for $J > J_c$. We say that *symmetry breaking* occurs at $J_c = 0$, where two new states emerge. In terms of information, we can say that the original information is doubled because two new attractors become available (Haken, 1977). These attractors can be visualized as the minima of a free-energy $\Phi(m, h)$ of the system, i.e.

$$\frac{\partial \Phi(m^*, h)}{\partial m^*} = 0$$

One possible free-energy function (also called Ginzburg-Landau potential) compatible with the previous equation for $m(t)$ is:

$$\frac{1}{N}\Phi(m,h) \equiv \phi(m,h) = \frac{\beta\rho}{2}m^2 + \rho \ln \left\{ \cosh\Big[\beta(\rho m + h)\Big] \right\}$$

in such a way that our dynamics is defined by a gradient system, i.e.:

$$\Gamma \frac{dm}{dt} = -\frac{\partial \Phi(m,h)}{\partial m}$$

and so a relaxation towards m_+^* or m_-^* (depending on the initial state $m(0)$) will occur.

We have introduced this particular case because it illustrates the early suggestions by Haken (1977,1988) about information processing in complex systems. To store information, the system has to be able to stabilize the attractors in deep enough minima (i. e. those defined by $\Phi(m,h)$) . But in order to process information, switching among attractors is necessary. If, through some self-regulated mechanism, a switch among attractors is available, processing becomes possible. Though it has been suggested that complex computation takes place in systems poised at critical points, here we suggest a different strategy. The system can store information by means of attractors and switch among them by moving *through* critical points.

By an adequate choice of the connectivity matrix, we can reach a particular global state by means of local interactions. Assuming that $J > J_c$ then the matrix

$$\Lambda_F = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

which defines the *ferromagnetic* FNN, is in fact the discrete counterpart of the previous mean-field model. For $\beta = 0$, half of the automata will be in state $+1$ and half in -1. When $\beta > 0$, symmetry breaking takes place. One of the possible attractors $\{m_+^*, m_-^*\}$ will be choosen.

Another possibility in order to reach a particular attractor can be to provide an external field $h > 0$. Then an asymmetry is introduced from the beginning and the system will be in one of the global states m_+^* or m_-^*. But there is an additional possibility. We take $\beta > 0$ and an appropiate Λ matrix such that transitions towards a particular state are more likely to occur. As an example, we can take:

$$\Lambda_0 = \begin{pmatrix} 1 & 1-\epsilon \\ 1 & 1-\epsilon \end{pmatrix}$$

(where $\epsilon \in (0,1)$) is such that leads to a FNN where almost all elements move towards the state $S_i = +1$, and so $m \to m_+^*$. In this way, we can *store* information.

Now the basic problem in order to perform computation is to have a mechanism of switching among attractors. In this way, we *process* information. One possibility, as stressed by Haken (1977,1988), is to be able to switch by means of fluctuations, which are larger as we approach $J = J_c$, the critical point where fluctuations diverge. But in this state the reliability of the system is lowered precisely due to fluctuations. So what we need is an additional mechanism which should be modified by the own system (say the agents), and able to reach different attractors under suitable inputs.

The previous mechanism of communication, which involves direct contact among neighbors, is just part of the story. Chemical communication in ants play a very important role and will be also considered in our study. Some general comments are of interest:

(a) Ants interact among themselves both by direct contact and by means of chemical substances (Wilson, 1971).

(b) By means of both phenomena, the ant behavior can switch from the current state to another. In any case, the ant can respond to the chemical signal and reinforce it. As a consequence, there is a non-trivial global behavior sustained by means of the individual activity of ants, being this activity simultaneously modified by the chemical field. The individual ants and the global activity pattern are related in two directions: from top to bottom and from bottom to top (see chapter 1).

(c) Computation (according to the characterization given at the beginning of section 3.1) takes place collectively. The ant colony perceive external signals and monitors its environment, being then able to make decisions. These decisions result from different inputs which can be conflicting (two different food sources, enemies, etc). Local inputs must be processed collectively in order to obtain an adequate global output. As stated by Wilson: "(local information) is judged principally, and perhaps exclusively, by the "electorate" response of the colony through all-or-none "voting" by the individual ants" (Wilson, 1971).

(d) The ant colony needs to be flexible. Once a given information is detected, chemical communication (as trails) makes possible the global response. The stability of the chemical signal is thus relevant for colony behavior. But after a given source of information is gone, the colony has to be able to switch again towards the "normal" state, defined by means of some task distribution. Again by Wilson: "the level of accuracy (of chemical communication) has been arrived at a compromise betwen the utmost effort of the ant's chemosensory apparatus to follow trails accurately and, simultaneously, the need to reduce the quantity and increase the volatility of

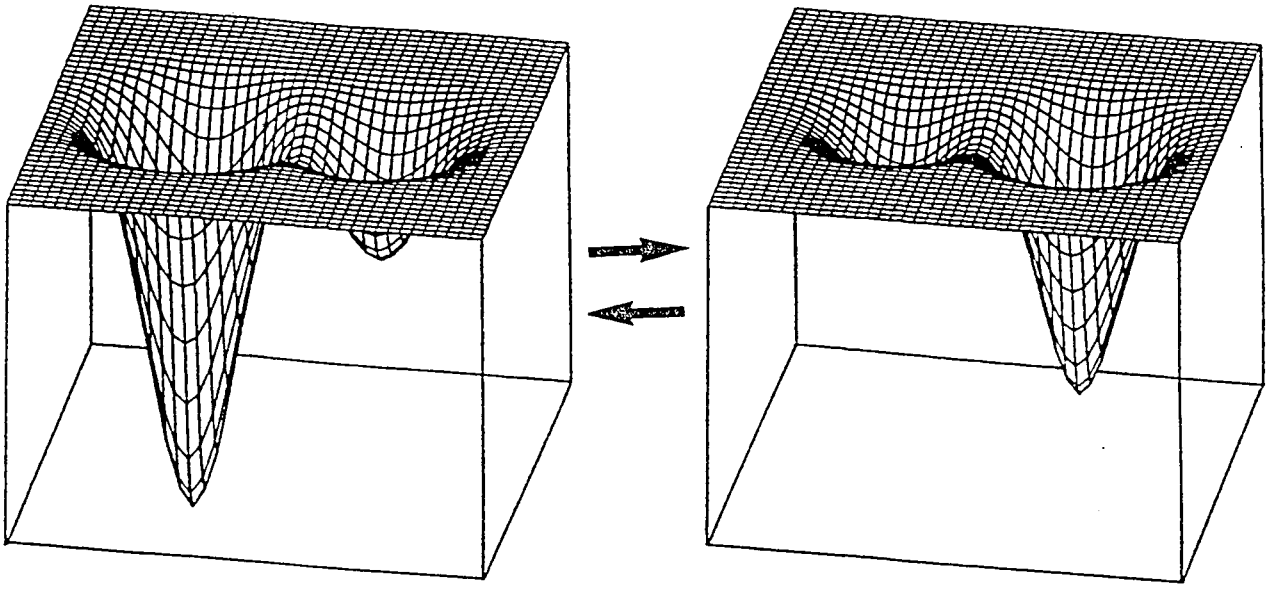


Figure 3.3: Switching between atractors. The numbers of individuals engaged in different tasks defines the global (colony) states. For each combination, an "energy" function is defined. The system spontaneously evolves towards those states characterized by the deepest valleys. As some external signal is introduced, the system can amplify it; by changing the landscape, a computational process can take place.

the trail substance in order to minimize overcompensation in the mass response" (Wilson, 1971)

The last two comments can be well represented by means of a simple diagram, which uses the idea of attractors, as shown in figure 3.3 (see also Gordon *et al.*, 1992). We can imagine the standard distribution of tasks as a given set of numbers of individuals engaged in different activities. This distribution is, in our approach , an "attractor" of the dynamical state. In figure 3.3a, it is shown by an energy landscape where the minimum represents the most probable distribution of states. Here α is the deepest valley, defining the most probable distribution, though may be other small valleys (β) could be present. As an external signal is detected, the system should be able to switch towards other attractors (i. e. $\alpha \rightarrow \beta$, by changing the landscape) in order to perform new tasks.

In our study, we consider a given chemical concentration $C(i,j)$ which can be detected and reinforced by the individuals. This substance, in absence of individuals, will have a simple dynamical evolution given by the diffusion equation (DE):

$$\frac{\partial C}{\partial t} = -\mu C + D\nabla^2 C$$

Here μ stands for the spontaneous rate of decay, D is the diffusion coefficient, and $\nabla^2 = \partial_x^2 + \partial_y^2$ is the two-dimensional Laplace operator. A discretization of the

previous DE will be used, where we will use the following numerical approximations:

$$\nabla^2 C(i,j,t) = \frac{1}{\delta x^2}\left(\sum_{k=-1,+1} C(i+k,j,t) + \sum_{k=-1,+1} C(i,j+k,t) - 4C(i,j,t)\right)$$

where we take $\delta x = 1$, and time is also discretized as:

$$\frac{\partial}{\partial t} C(i,j,t_k) = \frac{C(i,j,t_{k-1}) - C(i,j,t_k)}{\delta t}$$

where $1 \leq i,j \leq L$ and $t_k = k\delta t$ (here we use $\delta = 10^{-3}$). Clearly, our numerical calculation of dynamics of the chemical component defines two distinct time scales for individuals and the chemical field.

If individuals are present, a new term should be included in the previous PDE, involving the reinforcement of the chemical signal, as will be done below. It is interesting to mention here that chemical communication, together with tactile stimuli, is able to provide strong cohesiveness even to large societies. Army ants, which are among the largest societies, show spatially self-organized swarm raids that covers 1000 m^2 in a day (Deneubourg *et al.*, 1982). It is one of the best examples of collective decision making without centralized control of any kind. In this context, a remarkable study by Mikhailov (1993) on mass communication in distributed systems has been performed, involving the formal approach of neural networks together with chemical mediators.

Now we show how a NOR collective gate can be obtained in a simple way. First, we have to define the "normal" state of the network, which can be understood as a given distribution of tasks. For simplicity, we will define a colony state where the elements in our FNN are mainly in state $S_i = +1$. This can be achieved in several ways. Three situations are considered in the following subsections.

3.2.1 Matrix Λ_0, chemical switch

This is the simplest collective NOR gate, where a properly defined "collective" response Ω, the output of the gate, needs to be introduced. Here we take $\beta > 0$ and the connectivity matrix will be Λ_0. As a consequence of this particular choice, a given task distribution will be obtained, being the $S_i = +1$ individuals the most abundant.

Now let us define the two external "inputs", I_1 and I_2, necessary in order to construct the logical gate. Following our previous discussion, they can be two signals placed at two different points of the lattice (the environment). Here we use two opposite vertex. As a signal, we take a given fixed concentration C_0 which can

($I_i = 1$) or cannot ($I_i = 0$) be present. At a given time step (and for some time τ) we fix the state [2] of both points to C_0. Then if a given individual detects, at a given lattice point, a concentration $C > \theta$, where θ is a threshold, it reinforces the local concentration by an amount ψ. The local concentration acts on the local field in the following way:

$$h_i(t, C_i) = \sum_{j \in B_i(\rho)} J_{ij} S_j(t) - C_i$$

here C_i is the local concentration perceived by the single individual (located at a given lattice point). We see that in this way, the local field is modified, and the new transition probability makes possible to switch towards a different state. Physiologically, this is nothing but a change in the excitability of the individual automata. As C_i grows, it becomes more and more likely to switch towards $S_i = -1$. If the self-reinforcement of the chemical field is strong enough (and this will depend specially on μ and ρ) the *whole* system can switch. The local signals have been amplified and a new attractor has been created (as it was shown in figure 3.3). In figure 3.4 we show, for a particular set of parameters, the evolution of the three main quantities defined in our study: m, C and Ω. Here Ω is the discrete output obtained from our collective computation, and it is defined as

$$\Omega = \Theta \left[\sum_{j=1}^{N} S_j(t) \right]$$

where $\Theta(z)$ is the Heaviside step function.

We can see that after the appearance of the external signal, the system is able to switch towards the negative values and so $\Omega = 0$. In other cases (see below) the self-reinforcement of the chemical field is strong enough that the new state is maintained indefinitely. For very slow densities, both fluctuations and small information transfer makes the gate very unstable.

3.2.2 Matrix Λ_F, nonzero threshold

A different possibility is a FNN with ferromagnetic interactions (as defined by the matrix Λ_F) and where a given non-zero threshold ϕ is introduced. Again, individuals will move towards m_+^* (here we take $\beta > 0$). Now the local field is given by:

$$h_i(t, C_i) = \sum_{j \in B_i(\rho)} J_{ij} S_j(t) - \phi + C_i \tag{3.4}$$

[2] Many other possibilities are allowed, in such a way that the input signals are not coupled with the chemical field. For example, if an automaton detects an input $I_i = 1$, it leaves a given amount of chemical. These possibilities have also been explored, leading to similar results

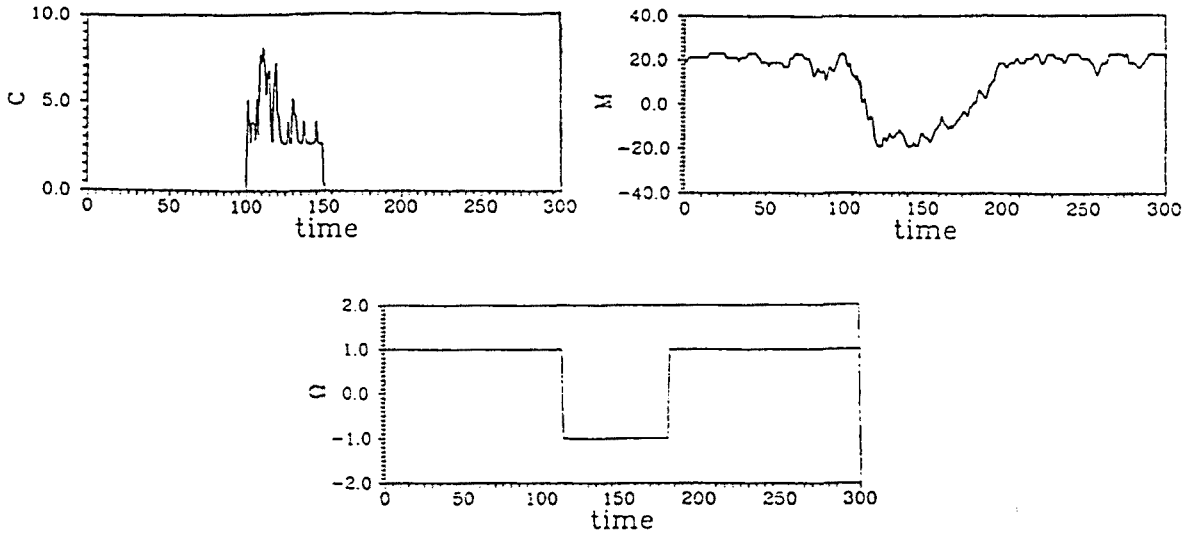


Figure 3.4: Collective NOR gate. An example of the dyamics of the three basic components of our model is shown (see text). As an external signal is introduced (from step 100 to 150) a chemical field is formed, and a switch is obtained. For this parameter combination, the switch is a transient situation. After some time steps, once the chemical field decay to zero, the previous state is recovered.

and the previous rules are the same as before. For an adequate choice parameters, a similar scenario is obtained. This can also be obtained using a different chemical field C_i^*, instead of a fixed threshold, ϕ. If no external input is present, the individuals can self-maintain this new field. When the external signal is detected, the new field would be more reinforced, and the transition can take place.

3.2.3 Matrix Λ_F, two chemicals

This is a more sophisticated situation. The system starts with a small $\beta > 0$ value and the Λ_F matrix. If no external inputs are present, both states $S_i = \pm 1$ are equally likely to occur. Now let us assume that two different types of inputs (which we can arbitrarilly call "0" and "1") are allowed, with concentrations $C^{(1)}$ and $C^{(2)}$. So now $I_i = 0$ means that $C^{(1)}$ is being introduced as an input, and $I_i = 1$ means that $C^{(2)}$ is used. Then two different chemical substances are specified as a symbol of the input alphabet. Both chemicals will have characteristic rates of decay (μ_1, μ_2) and diffusion rates (D_1, D_2).

Now the local field will read:

$$h_i(t, C_i^{(1)}, C_i^{(2)}) = \sum_{j \in B_i(\rho)} J_{ij} S_j(t) - C_i^{(1)} + C_i^{(2)}$$

where each chemical is acting, as we see, in opposite ways. Again, as automata find

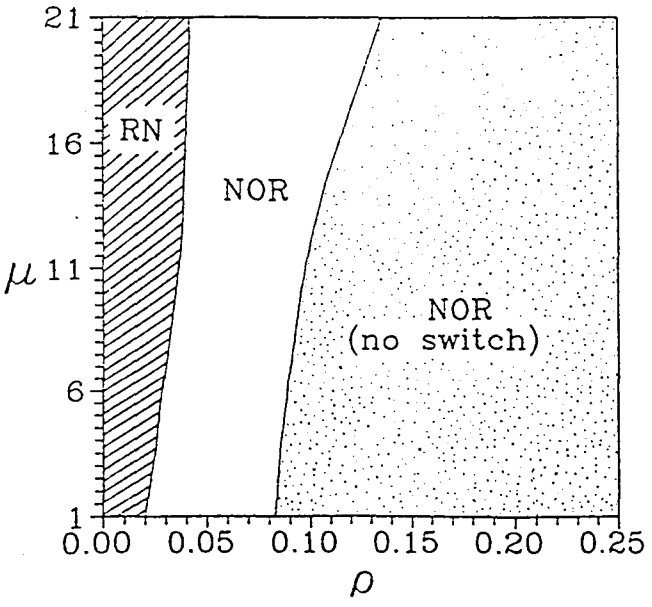


Figure 3.5: Phase space for the collective NOR gate (see text). Three qualitative domains are obtained. The dashed area (RN) is linked with those parameter combinations unable to define a NOR gate: random fluctuations are dominant. The dotted area is linked with a properly defined NOR gate, but the self-reinforcement of the chemical field is too strong (see text). The white area corresponds to a properly defined NOR gate, and the system returns to the starting attractor once external signals are removed.

locally any of the chemical signals, they reinforce the local values (if $C_i^{(k)} > \theta_i^{(k)}$, as before). The competition between both chemicals can result in different global outputs. If the specific rate constants are suitably chosen (for example, if $\mu_1 > \mu_2$ and $D_1 = D_2$) a NOR gate is obtained.

3.2.4 Parameter space

We have now to see how robust are these collective gates. Their behavior depends upon the parameters involved, and here we analyse the first type of gate. The following constants are used: $C_0 = 10$, $D = 0.2$, $\theta = 10^{-4}$ and $\psi = 0.5$. Using a $L = 20$ lattice, with $\epsilon = 0.35$ and $\beta = 2$, a detailed study of the parameter space (ρ, μ) has been performed. For each sample, we have discarded $T = 200$ transients in order to reach the colony attractor given by matrix Λ_0 and the external signals are introduced over a short period of time ($\tau = 50$ time steps). Then we study whether or not the colony is able to switch towards the new attractor and move back some time later. We take as correct those samples where the colony is able to switch and return to the initial distribution (here m_+^*) into a period of 150 time steps.

The phase space of our system, for the previous parameters, is shown in figure

3.5. Three regions are observed. The first one is the so called random network (RN, shaded area) obtained at very low densities of individuals, is a domain where the NOR gate cannot be built. Either random effects become dominant (and the FNN is switching randomly) or only if the two sources are present the system can switch. The dotted area shows a domain where the NOR gate works, but no turning back to the initial attractor is possible. The amplification of the incoming signals is too strong and becomes self-sustained. The white area is the most interesting: the NOR gate is built and the system is flexible enough to move back to the previous attractor. This happens at intermediate densities. As far as the collective gate can switch back to the previous state in a bounded interval of time, a characteristic time scale for computation can be defined. The main properties of this parameter space (as the separation in three well defined areas) are also obtained from the other implementations of the NOR gate.

3.3 Conclusion

In this chapter we have discussed the relation between computation and dynamical systems. According to a widely accepted characterization of *collective* computation (Forrest, 1990), this is determined mainly by the dynamical properties of the system and computation arises either because an external observer interprets the dynamical pattern as such or because the dynamical properties endow the system with functional properties. So then, what is usually done in the literature is to build a dynamical system that solves some sort of particular problem (Dorigo *et al.*, 1996). We have reviewed three well known examples of emergent or collective computation, according to the characterization mentioned above.

This framework is clearly inadequate if we want to connect collective dynamics with constructs from computation theory (Hopcroft & Ullman, 1979), such as finite automata or logical gates, which is, we think, indispensable in order to attain a *theory* of collective computation. A first attempt to deal with this problem has been done by means of an analysis of a particular problem involving the computational capacity of FNNs. As a starting point we have considered the possibility of building a NOR collective gate. A simplified situation was considered, where two external signals $\{I_1, I_2\}$ of some type are used, and a binary variable is assigned to them. A binary variable is also defined over the colony states in such a way that we can clearly construct a table for the gate.

The external signals are appropiately amplified by the automata, and the self-reinforced field acts on the individual individual states through a change in the

transition probabilities. By depending on the density of individuals and the rate of decay of the chemical field, the external inputs can be amplified. Then an emergent pattern is obtained: a self-sustained chemical field is created. As a consequence a global colony organization is reached. Some parameter combinations (ρ, μ) makes the system more or less flexible, eventually switching back towards other attractors as the external inputs are removed. Following this idea, we have obtained a NOR gate in several ways. One or two chemicals can be used. They can act on all types of individuals or in different ways for different states.

The existence of a domain of densities (ρ) where maximum flexibility is allowed is consistent with other experimental and theoretical works involving networks of patrolling ants (Adler & Gordon, 1992). The encounter rates among ants has been shown to be of extreme importance as an organizing factor for the colony behavior. In terms of dynamics, a low density makes possible to enhance fluctuations and switching among attractors. If the density is high enough, we can also ensure (at least transiently) the stability of the attractors. This compromise has also been observed in other situations (see chapter 2).

This work can be extended to more general situations. If an arbitrary set of chemical signals $\{C_i^{(k)}\}$ is involved (here $k = 1, 2, ..., f$) then the local field h_i perceived by the individual automaton will be:

$$h_i(t, \{C_i^{(k)}\}) = \sum_{j \in B_i(\rho)} J_{ij} S_j(t) + \sum_{k=1}^{f} \eta^{(k)} C_i^{(k)}$$

where $\eta^{(k)} \in \{-1, +1\}$ by depending on how each field behaves. The previous results would be then generalized. Equation 3.4 contains much of the ingredients of collective dynamics. Agent-to-agent interactions and the creation and interaction through chemical fields are both involved. The importance of each term will be different for different species, but their sensitivity to fluctuations and how the colony behavior is tuned are rather general problems. The appropiate transfer and processing of information requires parameter combinations not too far from phase transition points. In fact, recent theoretical studies shows that such points would play a prominent role in the evolution of social behavior, as early suggested by Wilson (Hölldobler & Wilson, 1990).

Two final, particular comments have to be considered for future studies:

(a) The FNN model is robust against noise. Though some deterministic cellular automata models (as the Game of Life) have been shown to be able to support universal computation (Berlekamp *et al.*, 1985) they are not robust when noise is present (as

it happens in natural conditions). A random change in a single automaton state can destroy the gate. A FNN finds the robustness through the amplification of incoming information and is only weakly dependent on failure of single units.

(b) The FNN is a spatially-distributed system. As a consequence, if different inputs enter to the FNN at different spatial locations, competition and pattern formation can occur. We can easily build up other types of collective gates in response to different types of inputs. It will be of interest to know how such -perhaps conflicting- inputs are processed into the system and how complex these computations would be. This also opens new problems. Though the NOR gate has been shown to be constructable, how several gates could be coupled in order to simulate more complex gates is far from trivial.

Thus, a first approach to the *general* problem of computation and collective systems has been introduced. There are some difficulties in going farther with this approach, such as the above mentioned simulation of more complex gates. The making of complex gates is likely to introduce ad-hoc mechanisms, so that the criticisms raised above to the theoretical work on computing with dynamical systems apply also to this case. Furthermore, a *non-uniform* model of computation, such as that of logical circuits (Balcázar *et al.*, vol. I, 1988), is not so desirable since a different circuit should be constructed not only for each problem, but for each length of problem instances. Thus, we have left the FNN formalism and we have looked for more abstract ways of dealing with this problem, in order to simplify it as much as we can. This is the subject of the next chapter.

Results in this chapter have been published in Solé & Delgado (1996) and in Delgado *et al.* (1997)

Chapter 4

Collective computation: Theoretical approach

Several ways of relating dynamical systems and computation have been mentioned in chapter 3, though any of them has been found adequate to build a theory of collective computation. The theoretical work on computing with dynamical systems (see for example Branicky, 1995) does not provide us with the tools with which to model real (either natural or artificial) systems and the "emergent computation" approach is too focused on solving particular problems to be useful for theoretical concerns. We have explored another way, that of building dynamical systems, not too far from those modelling real systems (the FNN in our case), that behave as simple computational devices (a NOR gate), but some difficulties have arisen that make unclear how to proceed from what we got. However, there is an alternative that is worth to explore, that of *computational mechanics* (Crutchfield, 1994b).

4.1 Computational mechanics

In section 3.1.1 we have reviewed three typical behaviors of ant colonies (collective decision making, collective sorting and task allocation) illustrating S. Forrest's characterization of emergent computation (Forrest, 1990). Let us go back to those examples and introduce computational mechanics with their help. First, let us rethink in terms of *information processing* what is going on when ants display the behaviors above mentioned:

(1) *Collective decision making*: The theoretical models reviewed so far may help to ascertain in which way collective food selection implies information manipulation. Both models capture one important feature of real ants: The creation of a field of pheromones that guides individual behaviour. The field (represented by σ^i in

Millonas' model and by $g(X_i)$, that is, trail recruitment in Beckers *et al.* model) plays the role of a long term memory. So, information storage is spatially distributed, which allows individual ants to manipulate this information, either reinforcing or not the field according to the local information available, that is, the "intensity" of the field on the surroundings of the ant. It is easy to see how information is transmited, it comes with the local changes of the field, though it can be transmited also from individual to individual in group/trail recruitment. Furthermore, we see that these are precisely the key aspects when modelling the dynamics of the phenomenon, so why do not ask for a quantification, in information processing terms, of the collective decision? We have seen, in the theoretical models, how abrupt changes -transitions- occur that allow the system to choose the best food, so what changes in the information processing mechanisms to get the correct behaviour? Of course there are not answers as yet to these questions, because there was not any adequate framework in which to pose them.

(2) *Collective sorting*: In this case we deal with a *stigmergic* process (see chapter 1) implying that information transmission is performed through the spatial distribution of A and B objects. This turns out to be also the memory used by the collective system to proceed with the clustering task, memory that is modified by individual activity through the picking up and putting down of objects. As the model suggests, the clustering process relies on some fluctuation in the random redistribution of objects that takes place in the initial steps, fluctuation that must be large enough to attract more objects of the same sort. Then, a change in an initially unstructured distribution occurs so that the clustering of equal objects can be initiated. This change in the information storage of our system (a change in structure), could it be quantified in computational terms?

(3) *Task Allocation*: Task allocation is a more general problem than the ones solved by the other reviewed emergent behaviours, though myrmecologists have speculated on some aspects of its information processing properties, for example information transmission: Robinson (1992) suggests that synchronized activity may enhance the sampling of information in individuals, assuming that information is transmitted from individual to individual, and Bonabeau et al. (1996) suggest that a "stimulus" associated with each task induces an individual to work in that task in a genetically determined way. Other information processing structures are far less clear. We could assume that the state of our system, in what concerns to which tasks are needed and which are not, is a sort of memory that induces individuals, by some unknown means (Robinson, 1992), to work out the currently needed task(s). Finally,

information transformation might be performed through task achievement, which changes the memory (state) of our collective system.

So, we have seen three examples of functional collective behavior in ant colonies in which either information storage, information transmission or information transformation plays an essential role. Furthermore, in all cases it is quite clear that the historical memory of the process must be considered at the level of the whole system, i.e. the pheromone field, the spatial distribution of objects and the tasks status (accomplished or not). If one wants to characterize the global state of the system the usual measures of structure in physical systems (Fourier spatial modes, wavelets and so on, see Hanson, 1993, chapter 1) may be applied. However *those measures do not provide the relevant information about the system*, because what the ant colony is really doing is to manipulate structured information, i.e. *computing*, in order to perform some activities essential for survival. This must be reflected, in some *measurable* way, in the collective state of the swarm and, of course, these measures must be computational measures. Thus, the analysis of physical systems by computational means, in order to uncover the *implicit manipulation of information embedded in these systems* (also called *intrinsic computation*), must link observations in real systems (either natural or artificial) with computational notions, which is precisely the purpose of the new field of computational mechanics.

There are some cases where we can find out systematically how information processing is being performed by the system of interest. Among these cases we may mention one dimensional cellular automata (Hanson, 1993; Mitchell et al., 1994), systems with quasiperiodic and period doubling route to chaos (Crutchfield & Young, 1989, 1990) and one dimensional spin systems (Crutchfield & Feldman, 1997). Let us illustrate how a computational structure may be embedded in a dynamical system by means of the ϵ-Machine Reconstruction Algorithm (ϵ-MRA), an algorithm with which to reconstruct Deterministic Finite Automata (DFA), that is, machines with a finite number of states (Hopcroft & Ullman, 1979) from a discrete symbol sequence. The dynamical system will be a discrete map: The logistic map (Wiggins, 1990) We will generate the sequence from the orbits of the logistic map (Crutchfield & Young, 1989, 1990)

$$x_{n+1} = 1 - \mu x_n^2 \tag{4.1}$$

First of all, our point of departure will be a discrete series of measurements from 4.1 and an arbitrary initial condition x_0. As we are measuring a continuous state, we will do so with an "instrument" with finite resolution ϵ, in our case we use the

partition

$$\Pi = \{x_n \in [-1, 0) \Rightarrow S_i = 0, x_n \in [0, 1] \Rightarrow S_i = 1\}$$

where $S_1 S_2 S_3 ...$ will be the sequence of bits $S_i \in \Sigma \equiv \{0, 1\}$ generated through the dynamics of the map 4.1. Π has not been chosen arbitrarily, since it can be shown that only a special class of partitions are suitable to study discrete maps. Indeed, if one wants to reach any conclusion with respect to the original system 4.1 studying only bit sequences, the partition must be *generating* or *Markov*, that is, there must be a finite-to-one correspondence between infinite sequences and initial conditions x_0 (Wiggins, 1990). We will be seeking a machine that is the best description of the series of measurements. We are, then, working with a *dynamical system plus instrument* (S, ϵ). Once we have a symbol string we have to detect "causal" states: The set of subsequences that renders the future conditionally independent of the past (a detailed discussion of what are the causal states is beyond the scope of this thesis, see the example below and Crutchfield, 1994b). These states and the probabilistic transitions among them (inferred also from the symbol sequence) will define the machine associated with the data stream.

Let us assume that the dynamical system 4.1 with the "instrument" Π generates the sequence of 21 bits

101010001000001000101

We parse the sequence with a sliding window of length D, $D = 5$ in this case, and create the parse tree of figure 4.1. Each node has associated an integer counting the number of times we have parsed a D-bit sequence whose path through the tree contains that node. Now we have to identify those nodes that are *future-equivalent*, that is, having a topologically identical subtree of some depth $L < D$. In this case, with $L = 2$, we can see in figure 4.2 how many different subtrees of depth 2 (also called *morphs*) there are in the tree of figure 4.1. Each node with a depth lower than Depth(tree)-L has a morph associated with it and can be labelled correspondingly as shown in figure 4.3. According to the numbering of nodes in figure 4.1 we have $A = \{1, 2, 4, 7\}$, $B = \{3, 10, 11, 8, 5\}$ and $C = \{6, 9\}$, that is, each label is a set of nodes.

Considering each label as a state, it is clear that we can build up the deterministic automaton of figure 4.4 from the labelled tree of figure 4.3, without transition probabilities. How do we find the probabilities? Let us come back to the node counters C_n, where n is the node. If a node n is a parent of a node n' with, say, a 0-edge joining the nodes, the probability

$$P(\text{label}(n) \xrightarrow{0} \text{label}(n')) = \frac{C_{n'}}{C_n}$$

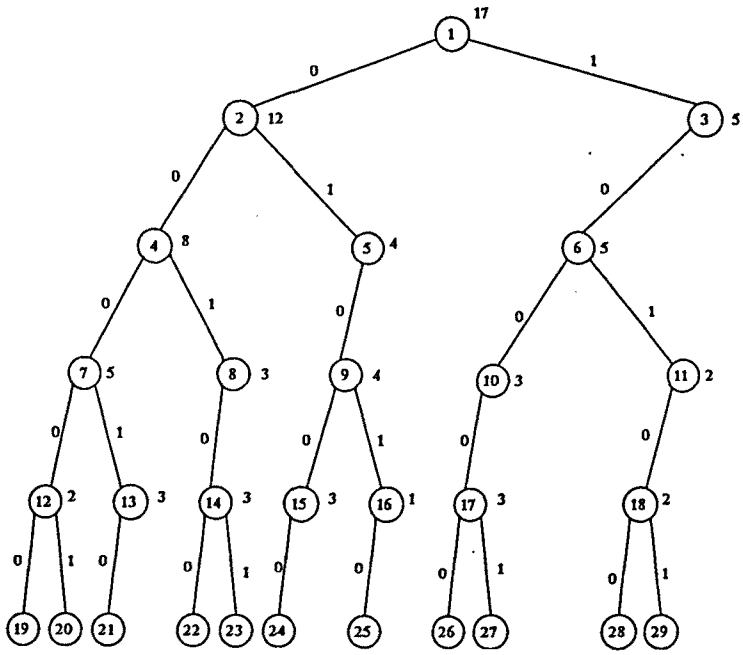


Figure 4.1: The result from parsing the bit string of the example detailed in the text with a sliding window of length 5 is the tree of the figure. Each node n is numbered and the counter C_n is shown aside.

so, according to figure 4.3, we have

$$P(C \xrightarrow{0} B) = P(6 \rightarrow 10) = 0.6$$

and

$$P(C \xrightarrow{1} B) = P(6 \rightarrow 11) = 0.4$$

But what happens with, for example, $P(A \xrightarrow{1} B)$? Is it equal to $P(1 \rightarrow 3)$?, or to $P(2 \rightarrow 5)$?. In case of ambiguities we will choose the transition highest up in the tree because the counts are larger and the statistics more reliable (Murphy, 1996), so $P(A \xrightarrow{1} B) = P(1 \rightarrow 3) \simeq 0.3$. The other probabilities are computed similarly, obtaining the automaton of figure 4.4. Figure 4.5 summarizes the entire process. There are some points that deserve further explanation. First, how do we choose the parameters D (window length) and L (morph depth)? If we choose D too large, there will be insufficient statistics, if it is too small it will fail to capture structure in the data. Likewise, if L is too small the models reconstructed, if any, would be inaccurate models of the data. In practice a series of reconstructions is performed for increasing D, choosing $L = \lfloor \frac{D-1}{2} \rfloor$, given a finite data set N. At first there will be different machines, due to inaccurate reconstruction because of D small. Afterwards there will be a period where the machine will be the same for

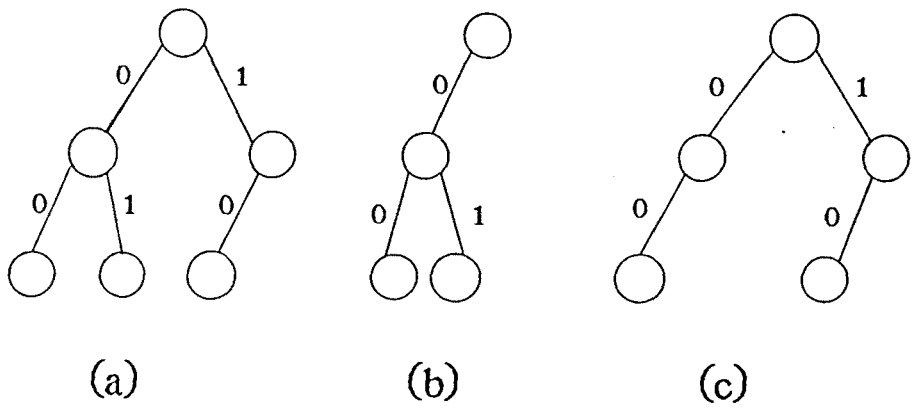


Figure 4.2: In the tree of figure 4.1 we find only these three morphs of depth $L = 2$: (a) under the nodes 1,2,4 and 7; (b) under the nodes 3,10,11,8 and 5 and (c) under the nodes 6 and 9.

increasing D. We will say that reconstruction is *stabilised.* For large D insufficient statistics appear and the reconstruction fails again. Second, is the result of the reconstruction always a Deterministic Finite Automaton (DFA)? The answer is no, we obtain sometimes a Nondeterministic (NFA) one. The reasons are diverse. In general, to assure DFA reconstruction some conditions have to be fulfilled: N large enough to have good statistics, the data stream has to be stationary and, of course, it has to be generated by a finite-state process. If any of these conditions is violated the reconstruction procedure never stabilizes. The more interesting case is that of infinite-state processes, where the reconstruction fails, though a close examination of the NFA could lead to a stack automaton (Hopcroft & Ullman, 1979), a higher level in Chomsky hierarchy (Crutchfield & Young, 1990).

From the ϵ-machine a measure of how complex is the automaton can be defined

$$C_\mu = -\sum_{\{S_i\}} P(S_i) \log_2 P(S_i)$$

where S_i are the *recurrent* states of the ϵ-machine and $P(S_i)$ is the stationary probability distribution of the ϵ-machine if viewed as a Markov chain (Luenberger, 1979). C_μ is called the *Statistical complexity.* Also, from the discrete symbol sequence (assuming now that is infinite) we can define the *entropy density* as

$$h_\mu \equiv \lim_{L\to\infty} \frac{H(L)}{L}$$

where $H(L)$ is the Shannon entropy of the distribution of length L subsequences

$$H(L) = -\sum_{\{S^L\}} P(S^L) \log_2 P(S^L)$$

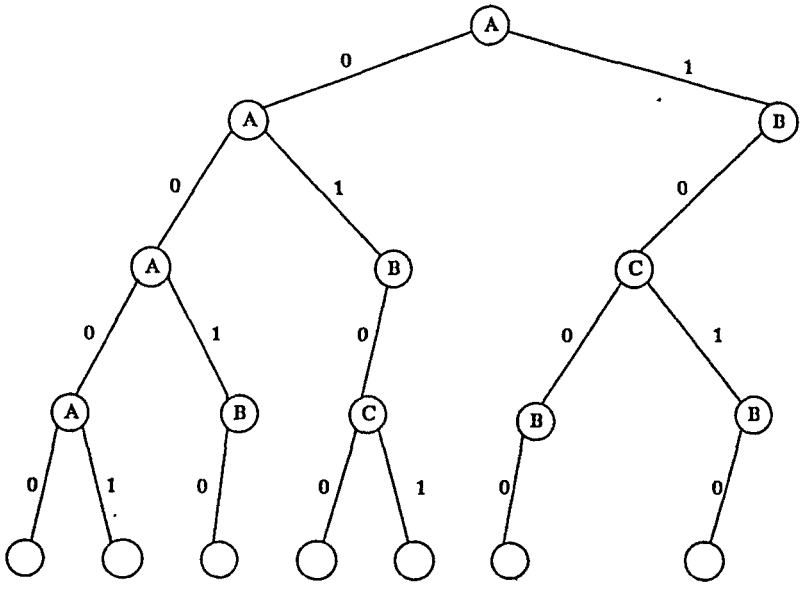


Figure 4.3: Nodes can be labelled according to the morphs of figure 4.2. Of course nodes with depth > 3 do not have any label. Nodes with the same label will represent the same state in order to construct the automaton (see text)

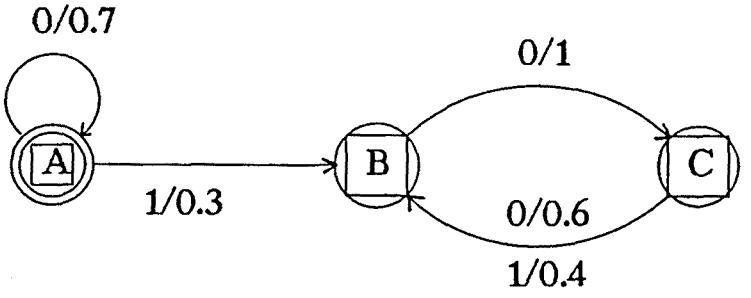


Figure 4.4: The DFA with probabilistic labels result of processing the 21 bit string shown in the text with the ϵ-MRA (see text).

Roughly, h_μ is a measure of how random is the sequence (for a complete discussion on h_μ, Crutchfield & Young, 1990 and Beck & Schlögl, 1993).

The ϵ-MRA is just a practical way of obtaining an intrinsic computation estimation. There is much more on computational mechanics; for example, in the case of one dimensional spin systems there is no need of using the ϵ-MRA, since C_μ can be found analitically (Crutchfield & Feldman, 1997). However, what we have reviewed up to now will be enough for our purposes (for a broad view of computational mechanics see Crutchfield, 1994b; see also Crutchfield, 1994a; Crutchfield & Young, 1989, 1990; Crutchfield & Feldman, 1997; Mitchell *et al.*, 1994; Hanson, 1993).

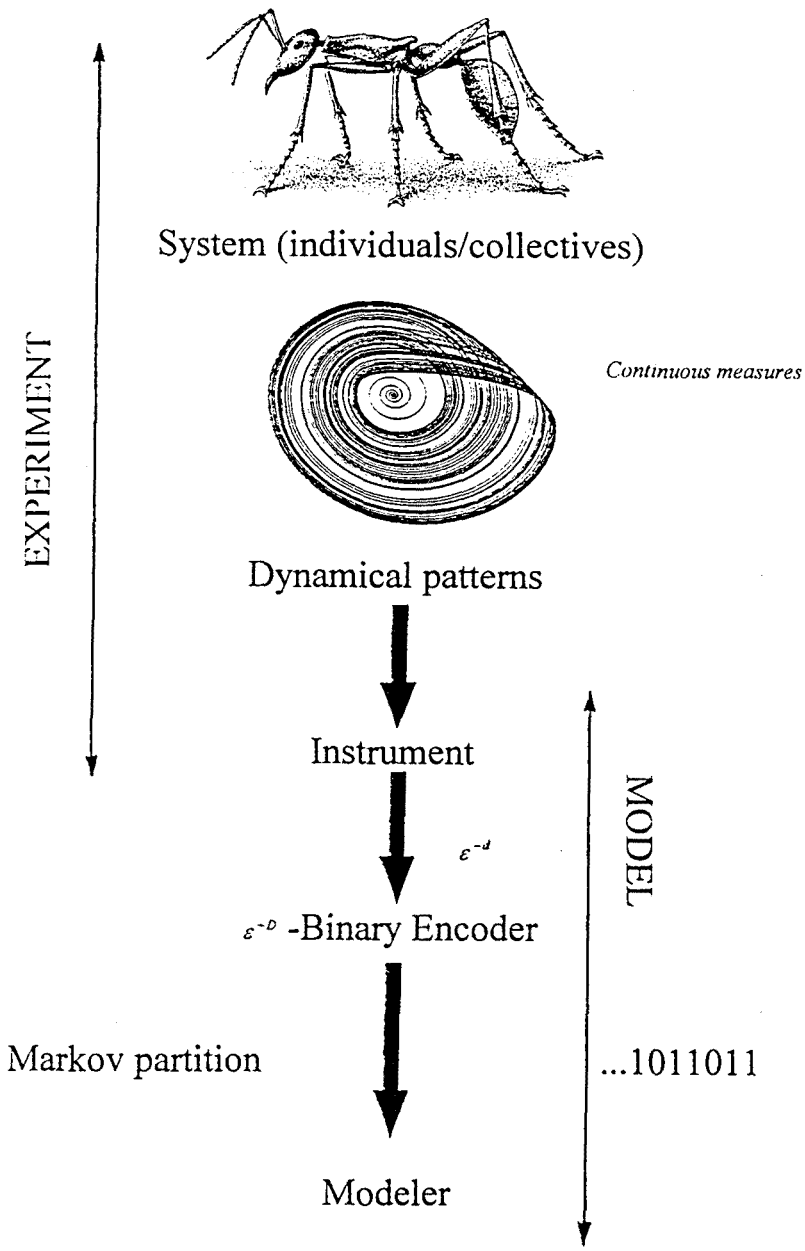


Figure 4.5: From the (possibly) continuous measures we take on the real system (either natural or artificial) to the discrete computational model we need to assume an observer with a finite resolution instrument. In the case of the logistic map, a Markov partition Π is the "instrument" with which the observer "filters" the measures.

4.2 RDAI and dynamical systems

Recently, researchers in RDAI have suggested that dynamical systems theory may be the appropriate framework for the development of autonomous agents. To cite a few examples, Beer (1995) shows how to approach the *synthesis* and *analysis* of autonomous agents problem with the theoretical apparatus of dynamical systems, building a six-legged walking agent (see also Beer, 1990); Sugawara & Sano (1997) study with dynamical systems the effectiveness of cooperative behaviour in groups of robots; Steels (1991) shows that the formal structure underlying systems with *emergent functionality* (a variation on the characterization of emergent computation by Forrest given in chap. 3) is a dynamical system; finally, Smithers (1994) makes a good case in favour of a dynamical systems approach to adaptive behaviour and cognition. This will be our working hypothesis to apply computational mechanics to the study of collectives of simple agents: *agents are dynamical systems and collectives of agents are collectives of dynamical systems.*

Furthermore, an artificial agent is built with some purpose in mind, it must have some *useful* behaviour that is of interest (to move objects, to follow paths and so on...). Identifying agents with dynamical systems allow us to ascertain their intrinsic computation, as we have seen above, but, what about their usable computation? that is, we are interested mainly in the capabilities of the agent to do something that we want it to do, so, how are these capabilities related with the intrinsic computation measures we are able to perform? Unfortunately, this question has not a definitive answer, though there is what we call the *Crutchfield's conjecture* (CC, Crutchfield, 1994a): Let us assume an agent A (dynamical system observed with the instrument I) with some ability to perform useful computations, ability that can be quantified in terms of its computational features, that is, the agent must behave with the computational power of a certain machine to solve the useful task; let us call this computational power $C^A_{us}(I)$; on the other hand, as a dynamical system it has a certain intrinsic computation associated, let us call it $C^A_{in}(I)$. Then, CC states that

$$C^A_{us}(I) \leq C^A_{in}(I) \tag{4.2}$$

that is,

> *intrinsic computation places an upper bound on what computations a dynamical system can support* [1]

[1]This conjecture was established for the first time in Crutchfield (1994a). As far as we know nobody has worked on its validity. Thus, it remains as an interesting and unexplored line of research

Now that we have identified (simple) agents with dynamical systems and we assume, from the CC, that intrinsic computation is an interesting measure *from an RDAI point of view* (since it places upper bounds on the useful computation the agents may perform), there is a final step in our reasoning to get a complete picture of how to deal with a theory of collectives of simple agents: what kind of dynamical systems do we choose to define the collective system? Of course there is not an absolute answer, since "dynamical systems" is a huge class of systems and a plethora of them might be chosen to work with. It is quite likely that the answer will depend on the problem we are focusing on. Our choice has been, we think, quite reasonable. We have decided to work with the logistic map 4.1 since: *a)* it is representative of a large class of dynamical systems (those with one maximum, see Beck & Schlögl, 1993), *b)* it displays, depending on the parameter μ, several dynamical behaviors such as fixed points, periodicity and chaos (Wiggins, 1990; Beck & Schlögl, 1993), *c)* it is a well known system and, in particular, its intrinsic computation has already been studied (Crutchfield & Young, 1989, 1990). However, see next section for a more detailed discussion of these three points.

So then, our theoretical work on collective computation will be performed on a collectives of logistic maps, also called *globally coupled maps* (detailed in the next section). Our initial work has been focused on a problem often stated (at least in the social insects community, see Moritz & Southwick, 1992) but, as far as we know, never formalised. Furthermore, we think it is a *general* problem in collectives of simple agents, so also of interest to the RDAI community. The problem is the following: Is there any relation between the "individual complexity" and the "collective complexity"? Could it be that collective "complex" behaviours are attained only with "simple" individuals?

4.3 Globally Coupled Maps

Because of the apparent arbitrarieness of our choice, we will first proceed to review some properties of the logistic map (the quadratic map family) before to proceed to the analysis of coupled systems.

4.3.1 Universality in Discrete Dynamical Systems

Now we will consider a brief review of some universal properties of discrete (-time) dynamical systems (DDS), with special attention to the family of the so-called one-dimensional uniparametric quadratic maps (UQM). The importance of this discussion is twofold. First, an overview of these models will be relevant in our discussion.

Second, such maps involve a surprising set of generic properties shared by a wide number of natural (chaotic) systems, from lasers and fluids to the heartbeat. Such universal features are important if we want to extend our conclusions to a relevant range of natural phenomena.

A DDS belonging to the UQM family is characterized by the following set of mathematical properties. Let us consider a map

$$x_{n+1} = \phi_\mu(x_n)$$

where $\phi_\mu(x) \in C^1([a,b])$. This one-dimensional (1-D) dynamical system belongs to the UQM class iff

- 1. $\exists x_m$ such that $\phi''_\mu(x_m) \neq 0$
- 2. $\phi_\mu(x)$ is monotonous in $[a, x_m)$ and $(x_m, b]$
- 3. The Schwarz derivative $\mathcal{S}(\phi_\mu)$ of ϕ_μ is negative, i.e.

$$\mathcal{S}(\phi_\mu) = \frac{\phi'''_\mu(x)}{\phi'_\mu(x)} - \frac{3}{2}\left(\frac{\phi''_\mu(x)}{\phi'_\mu(x)}\right)^2 < 0$$

Roughly, the UQM class is formed by a set of smooth functions with a single maximum in their interval of definition $[a,b]$.

The most interesting and relevant property of this family of maps is that they exhibit a set of bifurcations (as μ is changed) which shows generic properties. More precisely, as μ is increased from low to high values, starting from solutions of $x^* = \phi_\mu(x^*)$ displaying only fixed points, a *Feigenbaum scenario* (FS) is obtained for all $\phi_\mu \in$ UQM. The FS for the logistic map $x_{n+1} = \mu x_n(1 - x_n)$ is shown in figure 4.6. Here, the stationary orbits are shown after transients are discarded. We can see how the map shows period-doubling bifurcations at given critical values $\mu_1, \mu_2, \ldots$ where the number of points belonging to a p-periodic orbit doubles to $2p$. These periodic orbits $O_\mu^{(p)} = \{x_1^*, x_2^*, \ldots, x_p^*\}$ are such that

$$O_\mu^{(p)} = \left\{x_j^* | x_j^* = \phi_\mu^p(x_j^*), \forall j = 1, 2, \ldots, p\right\}$$

But it can be shown that an accumulation point μ_∞ exists for this sequence of bifurcations, leading to the emergence of a new type of dynamical regime known as *deterministic chaos*. The FS can be characterized by some universal constants, the best known of which is the δ-constant

$$\delta = \lim_{k\to\infty} \frac{\mu_{k+1} - \mu_k}{\mu_{k+2} - \mu_{k+1}} \approx 4.6692\ldots$$

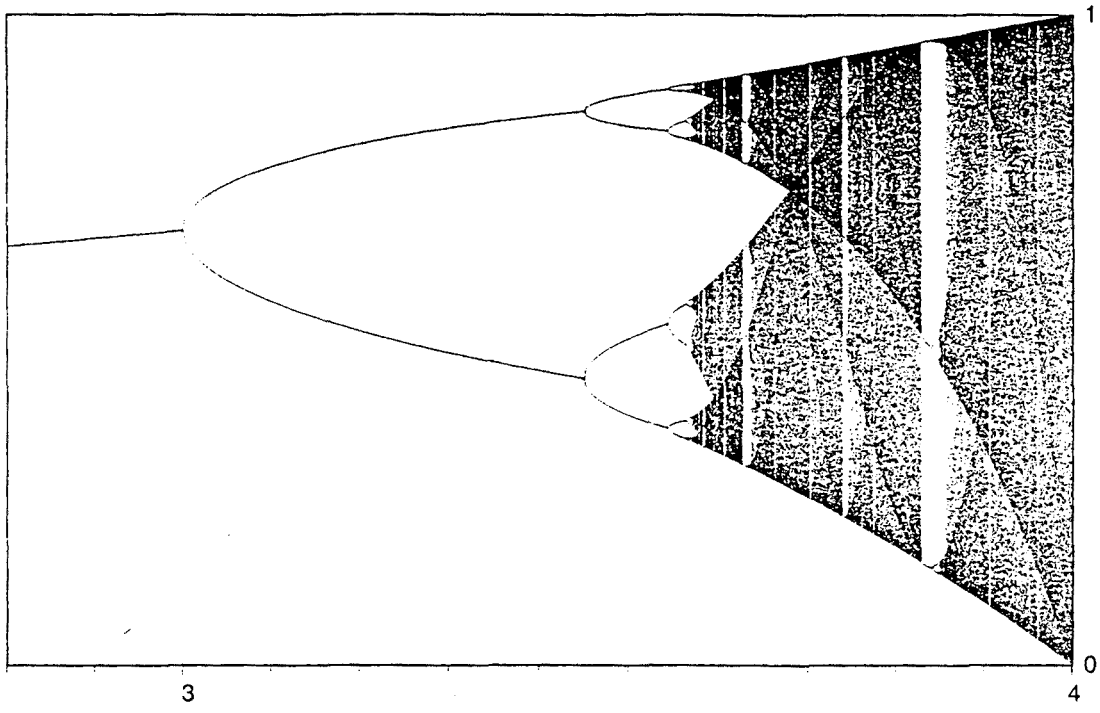


Figure 4.6: Period-doubling route to chaos. A sequence of bifurcations in the steady state of the logistic map is clearly seen as μ grows from 2.75. Beyond μ_∞, as μ keeps growing the system shows a complex sequence of periodic and chaotic dynamics.

and in fact a scaling relation links the critical point μ_∞ with the bifurcation points of the Feigenbaum cascade, through δ

$$\mu_n \simeq \mu_\infty - c\delta^{-n}$$

being c a constant. Beyond μ_∞ chaotic and periodic orbits are merged into a very complex parameter set structure.

The mention of μ_∞ in terms of critical points is not a metaphor. Following the general framework that has been used in this thesis, we should mention that the order-chaos transition shown by the UQM class is in fact completely analogous to a second order critical phase transition (Schuster, 1988).

Let us only mention a quantitative characterization of the degree of chaoticity which is in fact an order parameter of the transition to chaos: The Lyapunov exponent (LE). For a given $\phi_\mu(x) \in UQM$, the LE is defined as

$$\lambda_L = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \log |\phi'_\mu(x_k)| \tag{4.3}$$

which is $\lambda_L \leq 0$ for steady or periodic orbits, and $\lambda_L > 0$ for chaotic dynamics. If we take $|\lambda_L|$ as an order parameter close to μ_∞, it behaves as if in a critical transition.

It is not by chance that Feigenbaum solved the general problem of how UQM reach the μ_∞ point by means of the renormalization group method, which is one of the best known tools of the physics of critical phenomena (Binney *et al.*, 1992).

Although this family of one-dimensional maps may seem a poor description of natural systems, it became soon clear that, because of their universal behavior, they are indeed a close picture of many different systems displaying bifurcations and their transition to chaos. Fluid turbulence, laser beams, light emission by stars, chemical reactions, brain waves or the heartbeat are just some instances of natural systems where the Feigenbaum scenario has been detected (Schuster, 1988). What is more important, most of these systems, which are typically described by sets of $n > 3$ continuous differential equations, can in fact be reduced to one-dimensional maps by means of the so-called Poincare map transformation (Peitgen *et al.*, 1992). This is a crucial observation: It tells us that, essentially, all these systems are also members of the universality class of UQMs.

4.3.2 Coupled maps

GCM are usually defined by a set of nonlinear discrete equations (Kaneko, 1990a):

$$x_{n+1}(i) = (1-\epsilon) f_\mu(x_n(i)) + \frac{\epsilon}{N} \sum_{j=1}^{N} f_\mu(x_n(j))$$

where n is a discrete time step and $i = 1, ..., N$. The function $f_\mu(x)$ is defined by the logistic map 4.1 which is known to have a period-doubling route to chaos. GCM are in fact the simplest approach to a wide class of nonlinear networks, from neural networks to the immune system (Kaneko, 1990a). They have been shown to have remarkably rich behavior, partly similar to the mean-field model for the spin glass by Sherrington and Kirpatrick. Their behavior in phase space is very rich, showing clustering among maps. These clusters are formed by sets of elements with the same phase.

Before to explore the phase diagram of the Globally Coupled Maps (GCMs), we can show that this simple dynamical system is indeed a good (mean field) description of a wide variety of real networks.

GCMs are well known in spatiotemporal chaos as a mean field extension of coupled map lattices (CMLs). CMLs are a DDS where couplings involve only nearest neighbors (local interactions). In one dimension we have

$$x_{n+1}(i) = (1-\epsilon)\phi_\mu(x_n(i)) + \frac{1}{2}\epsilon\left[\phi_\mu(x_n(i+1)) + \phi_\mu(x_n(i-1))\right]$$

where $i = 1, \ldots, N$. The dynamics described by GCMs consist in fact of a parallel nonlinear transformation and a feedback from the mean field. Here the mean field follows from the relation

$$\sum_i x_{n+1}(i) = \sum_i \phi_\mu(x_n(i))$$

In the short-ranged CMLs a rich variety of pattern dynamics and phase transitions are known to occur. And it can be shown that GCMs give a mean field theory for such a rich variety of phases in the pattern dynamics of CMLs. We should mention here that CMLs share a considerable range of properties with cellular automata (Chaté & Manneville, 1989). This analogy reinforces the relevance of GCMs as models of discrete distributed systems.

Now let us consider a simple transformation

$$y_n(i) = \phi_\mu(x_n(i))$$

which we now introduce in the GCM

$$y_{n+1}(i) = \phi_\mu \left((1-\epsilon) y_n(i) + \frac{\epsilon}{N} \sum_{j=1}^{N} y_n(j) \right)$$

which is very close to a neural network model if $\phi_\mu(x)$ is a sigmoidal function such as $\tanh(\beta x)$. The connection between GCMs and relevant problems in neural network theory has been discovered by Kaneko (1990a). The physics of GCMs is not far from that of the most relevant complex disordered systems like spin glasses. But GCMs are also a very good approach to most network-like biological systems (Kaneko, 1994). Many biological systems display clustering of synchronization by chaotic instability, and the relevance of such clustering to ecological, immune, neural and cellular networks has been deeply analyzed with GCMs. In terms of computational processes, it has been shown that GCMs display clustering but also coding, switching, hierarchical ordering and control when a network of chaotic elements is used (Kaneko, 1990a). One of the implications of Kaneko's work, as pointed out by J. Casti, is in fact that the coherent structures emergent from the connection of chaotic elements (as we will see below) may be exactly what is needed to account for things like the persistence of neural memory in a disordered neural network (as it probably happens in the brain) or, more generally, the emergence of patterns from a collection of disordered individual agents (Casti, 1992). This fact and their universal properties, makes GCMs a perfect candidate to explore the interplay between individual and collective properties in terms of computation in dynamical systems.

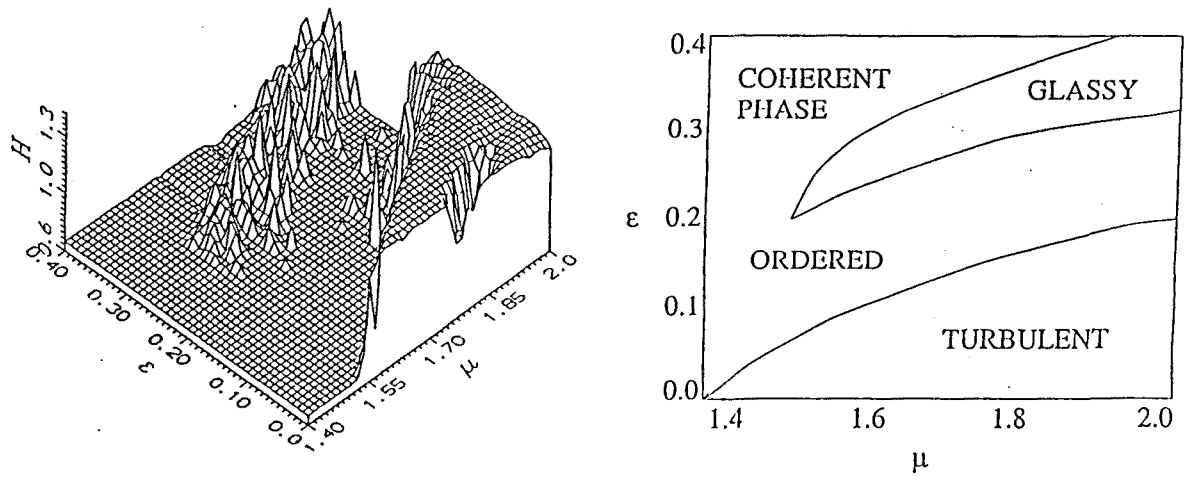


Figure 4.7: Information-Theoretic measures are able to discriminate among the different phases of GCM dynamical behaviour. *Right:* Joint Entropy for $0 \leq \epsilon \leq 0.4$, $1.4 \leq \mu \leq 2.0$ and $N = 100$. *Left:* GCM phase space (after Kaneko, 1990a). The joint entropy is largest at the turbulent phase when all the binary pairs are equally explored. It is ln(2) for the ordered and coherent phase and it takes intermediate values at the glassy phase. ϵ and μ are dimensionless parameters of the GCM

The phase space of GCM exhibits several transitions among coherent, ordered, intermittent and turbulent phases. These phases are well characterized in terms of the so called cluster distribution function $Q(k)$ (Kaneko, 1990a) and can also be well characterized, as shown in this section, by means of information-theoretic measures (Ash, 1965).

In each phase, a given number of clusters N_r involving r maps will be observed. Specifically a cluster is defined by the set of maps such that $x_n(i) = x_n(j)$, for all maps belonging to the cluster. We can calculate the number of clusters of size r, and for a given phase we have a set $\{N_1, N_2, ..., N_k\}$ of integer numbers. Then the $Q(k)$ function is defined as the fraction of initial conditions which collapse into a given $k-$cluster attractor (i. e. the volume of the attraction basin). An additional useful measure will be the mean number of clusters, R_μ, defined as $R_\mu = \sum_k kQ(k)$.

Here we also consider an information-based characterization of the different phases by means of the sequence of bits $S_i^j \in \Sigma \equiv \{0, 1\}$ generated through the dynamics of the i-th map, under the partition Π. We can compute the Boltzmann entropy for each map,

$$H^i(\Sigma) = -\sum_{S_i^j=0,1} P(S_i^j) \log_2 P(S_i^j)$$

and the joint entropy for each pair of maps,

$$H^{il}(\Sigma) = -\sum_{S_i^j=0,1}\sum_{S_l^r=0,1} P(S_i^j, S_l^r)\log_2 P(S_i^j, S_l^r)$$

From the previous quantities, we can compute the information transfer between two given units. It will be given by:

$$M^{il}(\Sigma) = H^i(\Sigma) + H^l(\Sigma) - H^{il}(\Sigma)$$

These quantities have been widely used in the characterization of macroscopic properties of complex systems modelled by cellular automata and fluid neural networks (chapter 2). As a way of quantifying complexity, it has been shown that information transfer is an appropiate measure of correlations and in this context it is maximum near critical points (Solé *et al.*, 1996). Because our interest is in the computational structure behind the observed dynamics, we expect to have some well defined relations between computational complexity and information transfer. Using these measures (see figure 4.7), the four basic phases exhibited by GCM are:

(1)*Coherent phase*: the system is totally synchronous, i. e. $x(i) = x(j)$ for all i, j. The motion is then described by a single map $x_{n+1} = f_\mu(x_n)$ and the stability of this single attractor can be analyticaly characterized (Kaneko, 1990a). If λ_0 is the Lyapunov exponent for the single map, the Jacobi matrix is simply given by

$$\mathbf{J}_\mu = \partial f_\mu/\partial x_n\Big[\,(1-\epsilon)\,\mathbf{I} + \frac{\epsilon}{N}\,\mathbf{D}\,\Big]$$

where $\mathbf{I}$ and $\mathbf{D}$ are the identity matrix and a matrix of ones, respectively. From the Jacobi matrix we can get the following stability condition:

$$\lambda_0 + \ln(1-\epsilon) < 0$$

Here almost all basins of attraction are occupied by the coherent attractor and $Q(1) = 1$, so we have $R_\mu = 1$.

In terms of information transfer under the generating partition, we will have $H^i(\Sigma) = H^l(\Sigma)$ (both maps are visiting the same points) and $P(S_i^j, S_l^r) = \delta_{jr}/2$ so it is easy to see that in this phase we have $H^{il}(\Sigma) = H^i(\Sigma)$ and the mutual information is given by $M^{il} = H^i$. The information is totally defined by the entropy of the single maps, as far as the correlations are trivial.

(2)*Turbulent phase*: this corresponds to the other extreme in the dynamical phases of GCM. Here we have that the number of clusters are such that $R_\mu \approx$

N. A first look at the dynamics of single maps seem to suggest that they behave independently. Under this hypothesis, the entropies can be easily estimated. If the maps are independent, then we have again $H^i(\Sigma) = H^l(\Sigma)$ but the joint probabilities will be such that $P(S_i^j, S_l^r) = P(S_i^j)P(S_l^r)$ and so we have $H^{il}(\Sigma) = 2H^i(\Sigma)$ and as a consequence the mutual information will be zero. A close inspection of the numerical values for the mutual information shows, however, that $1 \gg M^{il} > 0$, so some amount of correlation is still present. Specifically, we found that typically $10^{-6} < M^{il} < 10^{-3}$. This result was obtained by Kaneko (1990b) in a remarkable work where it was shown that GCM violate the law of large numbers (LLM). This hidden order is shown to exist by means of the analysis of the local fields, defined as $h_n \equiv N^{-1} \sum_j f_\mu(x_n(j))$. The study of the mean square deviation (MSD) of this quantity, which is expected to decay as $O(1/N)$ if the units are really independent, was shown to saturate for a given $N \geq N_c(\mu)$. The analysis of the density distribution for two maps gives a pair of continuous functions $P_i(x)$ and $P_j(y)$ (i. e. $\int P_i(s)ds = 1$) and a joint distribution $P_{i,j}(x,y)$ (with $\int\int P_{i,j}(x,y)dxdy = 1$) which makes possible to define a continuous mutual information

$$M_{i,j} = -\int\int \log\left[\frac{P_{i,j}(x,y)}{P_i(x)P_j(y)}\right] dxdy$$

and, after averaging over space and time it also shows a saturation when N gets large. Numerical experiments gave $M_{i,j}(N \to \infty) = O(10^{-3})$, consistently with our bounds for the binary partition. Such remaining finite correlation is the origin of the breakdown of the LLN.

(3) *Ordered phase*: here we have a small number of clusters with many units. Specifically, we have $Q(k) = 0$ for $k > k_c$ (where k_c does not depend on N) and

$$Q_L(k) \equiv \sum_{k>N/2} Q(k) = 0$$

and $Q(1) \neq 1$. We also get $R_\mu = b \ll N$. Again, a large number of elements will share the same state, and we can easily estimate the entropies and information transfer. Given two maps, they could belong to the same cluster or to two different clusters. In the first case, we get the same result than in the coherent phase, and the same occurs if they belong to clusters which are in phase. If the maps belong to two clusters which are not in phase, we have $H^i(\Sigma) = H^l(\Sigma) = \ln(2)$ and now $P(S_i^j, S_l^r) = (1 - \delta_{jr})/2$ so again we get $M^{il} = H^i$, as in the coherent phase (see figure 4.7).

(4) *Glassy phase*: also called intermittent phase, in this domain of parameter space we have many clusters, but they have a wide distribution of sizes. We have $\sum_{k>N/2} Q(k) > 0$ and also $\sum_{k<N/2} Q(k) > 0$. So $R_\mu = rN$ with $r < 1$. Here the competition of some attractors with different cluster size leads to frustration (Kaneko, 1990a). Following our previous arguments it is not difficult to show that $0 < M^{il}(\Sigma) < \ln(2)$. So in this phase the joint entropy has a finite (but not large) value, as expected from the existence of a decaying distribution of cluster sizes.

So we have shown that the use of information-based measures involving the previously defined generating partition provides an accurate characterization of the GCM phases. As we can see, some phases have a high information transfer while others have a nearly zero correlation among units. The basic qualitative observation of this phase space is that the greater the nonlinearity (the parameter μ) the more widespread is the disorder and that the greater the averaging effect (parametrized by ϵ) the more the overall coherence. So each unit in the GCM is subject to two competing forces: the individual tendency to chaos and the tendency to conformity arising from the averaging effect of the system as a whole.

4.4 Individual vs. Collective

4.4.1 Problem statement

The topic we want to address is easy to state: *the more complex a society, the more simple the individual* (Klaus Jaffe, cited in Moritz & Southwick, 1992). This sentence, of course, concerns to social insects, among which we will take ants as a main example. It is a well known fact that all living species of ants are eusocial (i.e. all species have the following properties: cooperation in caring for the young, overlap of at least two generations capable of contributing to colony labor and reproductive division of labor, Hölldobler & Wilson, 1990), nevertheless there exist large differences among species, with respect to the number of ants that compose the colony, their collective capabilities and the cognitive skills of individuals. A specific example is that of recruitement strategies: there is a clear correlation between the size of the colony and the behavioral sophistication of individual members (Beckers *et al.*, 1990). In one extreme we find the more advanced evolutionary grade: mass communication (information that can be transmitted only from one group of individuals to another group of individuals, according to Hölldobler & Wilson, 1990, p. 271). Mass communication is the recruitement strategy used by Army Ants (e.g. *Eciton burchelli*), whose colonies are composed by a huge number of individuals, who are,

nevertheless, almost blind and extremely simple in behaviour when isolated. The other extreme is occupied by those ants using individual foraging strategies (e.g. the desert ant *Cataglyphis bicolor*), who displays very complex solitary behaviour.

Our interest here is not so much to study this remarkable feature of eusocial insects, as to see if this could be a general trait of collectives of agents. That is, *is there a trade-off between individual complexity and collective behaviour, in such a way that complex emergent properties cannot appear if individuals are too much complex?*

In order to go on with our work, let's start looking thoughtfully at the concept of emergence. According to Hermann Haken (1988), the emergent properties of a system can be studied with the notion of *order parameter* and its associated slaving principle. We can look for an answer in two directions: from the individual to the collective and vice versa. Immediately we can discard the former, because the simplest individuals are those who display collectively the most complex behaviour. So, we can ask now a more concrete question: *what kind of behaviour the collective induces on the otherwise simple individual to attain emergent functional capabilities?* Of course we can answer it from an evolutionary point of view, arguing that adaptation to the environment is the ultimate reason of those diverse features of ant colonies. This is not the unique answer we can provide (Goodwin, 1995), because we can also look for relations between the order parameter and the individuals in such a way that, perhaps, complex solitary behaviour imposes severe constraints on the behaviour that a collective would induce on individuals. This would be a structural solution of our problem, and it will be the answer we are seeking.

4.4.2 Collective-induced computation

Now, the collective system we are working on is a Globally Coupled Map and our individual will be a randomly chosen logistic map of the system. Our purpose is to see how the collective is (or is not) able to induce more complex behaviour than that the individual is able to show.

Complex Individuals

Given a logistic map (our individual) a high statistical complexity is observed for μ close to μ_∞, i.e. the onset of chaos. There we need a large number of states to model the high periodicity of the orbits. We have chosen $\mu = 1.4$ whose statistical complexity is $C_{1.4} \simeq 4$. As we can see in figure 4.8 *(a)*, this automaton has a large number of states. The next step is to define a GCM with $\mu = 1.4$, and look at the statistical complexity of an individual (all are in principle equal) chosen at random,

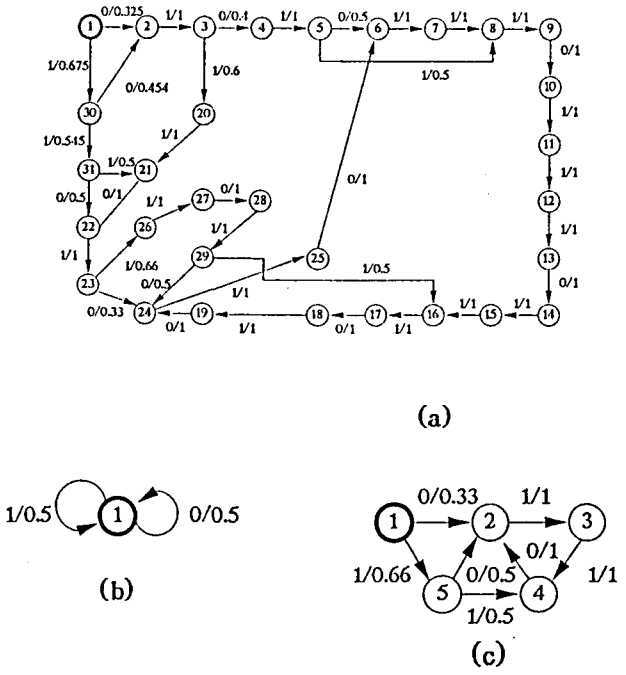


Figure 4.8: DFA with probabilistic labelings resulting from the ϵ-MRA applied to *(a)* logistic map with $\mu = 1.4$; *(b)* logistic map with $\mu = 2$; *(c)* logistic map with $\mu = 1.75$. These are the individuals over which we will check if the collective can induce more complex behaviour. As is obvious from the automata, *(a)* is much more complex than *(b)* and *(c)* (see text). In all cases the ϵ-MRA parameters are $M = 10^7$, $D = 32$ and $L = 16$. In *(a)*, *(b)* and *(c)* the state 1 is the initial state, all other states are accepting states. μ is the dimensionless parameter of the logistic map

say, i, as the degree of interaction increases, i.e. we examine $C^i_{1.4}$ as the parameter ϵ goes from 0 to 0.4.

The result is simply that there are no changes (as can be seen in figure 4.9). The intrinsic computation of the individual remains to be the same, $C^i_{1.4} \simeq 4$, no matter how large is the interaction with the rest of the system. So, the collective has not been able to induce any kind of added complexity to the individual. In this case there is no emergent behaviour. The collective behavior can be reduced to that of the individuals.

Simple Individuals

If we take $\mu = 2$ the logistic map has completely chaotic dynamics. It is, in statistical complexity terms, the same as a fair coin toss. So, its automaton has $C_2 = 0$ with just one state (figure 4.8 *(b)*). Now, we can apply the ϵ-MRA to the symbolic dynamics (i.e. the bit string of length M) of an individual chosen at random among the N that compose the GCM. The ϵ-MRA failed to reconstruct any automaton in the turbulent

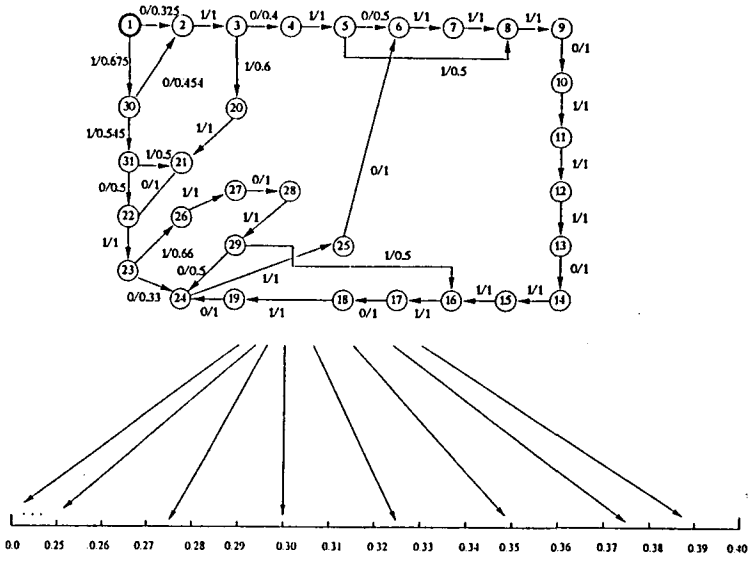


Figure 4.9: If we have a complex individual, no matter how much interaction it receives, its behaviour will not change. The collective cannot induce on the individual any kind of added behaviour. In the figure, the individual possesses the same statistical complexity, for all ϵ. Parameters of the ϵ-MRA: $M = 10^7$, $D = 32$ and $L = 16$. All the automata have 1 as initial state, and all other states are accepting states. ϵ is a dimensionless parameter of the GCM (see text).

phase (neither for $\mu = 2$ nor for $\mu = 1.75$, in the next subsection). This could be because of high dimensional chaos and the existence of supertransients (K. Kaneko, personal communication). In any case, it seems that the stationarity assumption were not fulfilled causing the non convergence of the ϵ-MRA (see Hanson, 1993 chap. 5). There are also some values of ϵ in the ordered and the glassy phase where no finite automaton was obtained. The reason here is the fine structure of those phases (K. Kaneko, personal communication). Our result is somewhat surprising (figure 4.10). If we exclude the automaton at $\epsilon = 0.26$ and the gaps at $\epsilon = 0.27$ and $\epsilon = 0.28$ (which indicates some kind of irregular behaviour in the regions, although according to the phase space of figure 4.7 we should have ordered behaviour) our individual reaches high complexity, $C_2^i \simeq 3$, near the boundary of the turbulent phase. Beyond this point we find the same automaton around $\epsilon \simeq 0.295$, perhaps pointing out another boundary (that of the above mentioned irregular behaviour). After that the complexity decreases with ϵ while going deeply into the ordered phase: first $C_2^i = 2$ at $\epsilon = 0.31$, then it goes down to $C_2^i = 1$ at $\epsilon = 0.32$, $\epsilon = 0.325$ and $\epsilon = 0.33$ to end up in $C_2^i = 0$ at $\epsilon = 0.34$ and $\epsilon = 0.35$. Complexity increases slightly again at the glassy phase: $C_2^i \simeq 1.585$ at $\epsilon = 0.375$ and $\epsilon = 0.39$. The more complex behaviour is displayed near phase boundaries, as has been observed also in other

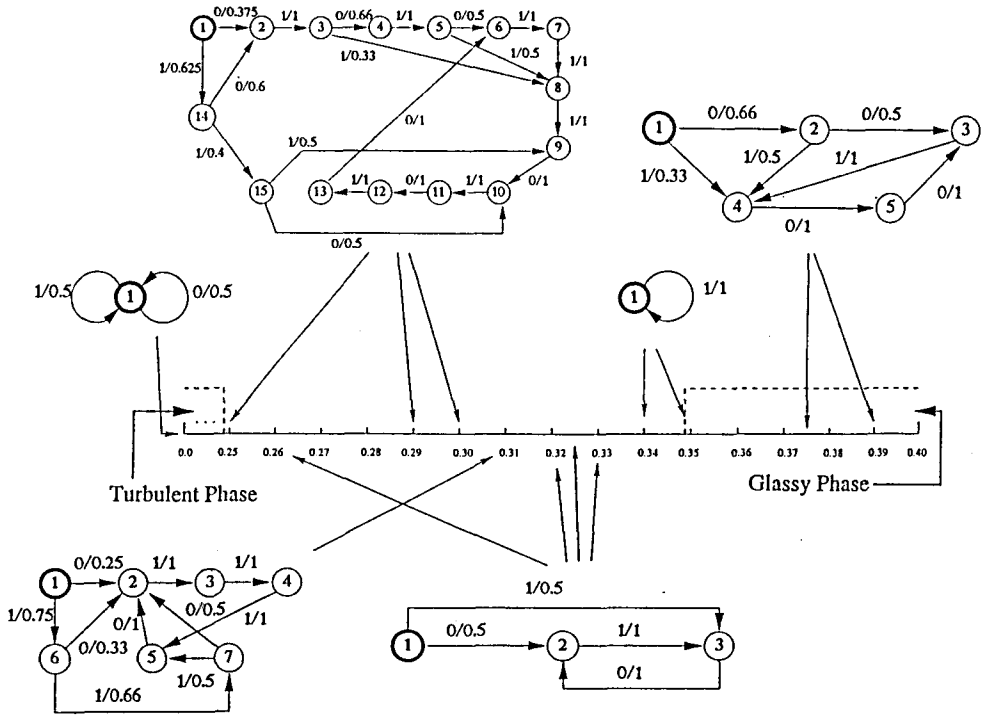


Figure 4.10: With a simple individual like that of $\epsilon = 0$ (in this figure), the collective is able to impose additional behaviour on the individual. We have a decreasing complexity from turbulent phase boundary onwards with increasing ϵ, except in the region of 0.27 (see text). We can observe also a slight increase in complexity at the Glassy phase. Parameters of the ϵ-MRA: $M = 10^7$, $D = 32$ and $L = 16$. All the automata have 1 as initial state, and all other states are accepting states. ϵ is a dimensionless parameter of the GCM.

systems (Solé *et al.*, 1996).

If we compare this case with the previous one, we see that simple individual behaviour allows the interaction to create more sophisticated behaviour in the individual, inducing a certain amount of statistical complexity that was not present at the individual level. So, a coordinated behaviour, which the individual is unable to show, emerges from the collective through interactions.

Intermediate Individuals

Here we have $\mu = 1.75$ with an individual of complexity $C_{1.75} \simeq 1.585$ (figure 4.8 *(c)*) and we take a logistic map randomly from a GCM with the same μ value. In this case, as in the previous one, we find maximum intrinsic computation at the boundary between the turbulent phase and the ordered phase. In fact, the automaton in this boundary is the same one we found at the same boundary for $\mu = 2$. Although the individual is more complex than that of $\mu = 2$ we can observe the same behaviour

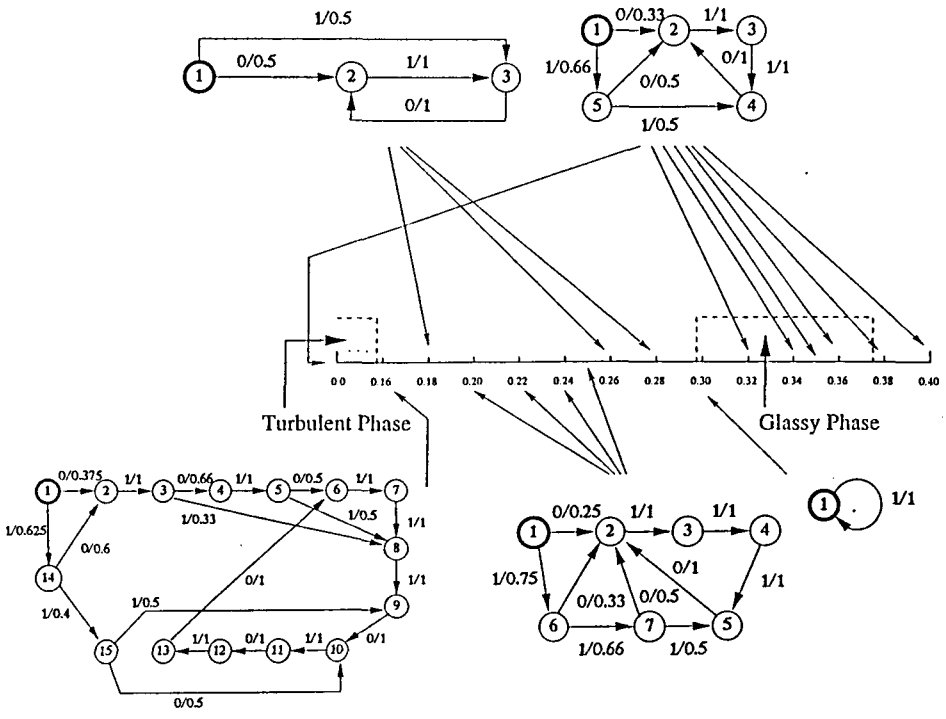


Figure 4.11: Here we see intermediate behaviour between the cases shown in figure 4.9 and figure 4.10 Just for $\epsilon = 0.16$ we found much greater complexity than that of the individual, and we have also bear in mind that the individual statistical complexity is $C_{1.75} \simeq 1.585$, so that the increment is not as large as in the $\mu = 2$ case (see text). Parameters of the ϵ-MRA: $M = 10^7$, $D = 32$ and $L = 16$. All the automata have 1 as initial state, and all other states are accepting states. ϵ and μ are the dimensionless parameters of the GCM.

of the automata with growing ϵ: at $\epsilon = 1.6$ we get a statistical complexity of $C^i_{1.75} \simeq 3$, at $\epsilon = 0.2$, $\epsilon = 0.22$, $\epsilon = 0.24$, $\epsilon = 0.25$ and $\epsilon = 0.26$ statistical complexity decreases to $C^i_{1.75} \simeq 2$, then statistical complexity keeps decreasing down to a value of $C^i_{1.75} \simeq 1$ ($\epsilon = 0.26$ and $\epsilon = 0.28$) and finally it reaches the zero value at the boundary of the glassy phase. However, this picture fails at $\epsilon = 1.8$, perhaps due to a small window located in the region of that ϵ. Again, at the glassy phase, there is a slight increase of complexity, i.e. $C^i_{1.75} \simeq 1.585$, that is precisely its individual value. The individual keeps this complexity value until $\epsilon = 0.4$, although there is another boundary, separating glassy and coherent phases. It is clear that now the individual is enough complex to have non zero statistical complexity and it is enough simple to let the collective to induce some amount of complexity. Of course the complexity growth is not as large as was in the previous case, because here the maximum complexity reached at the boundaries is the same that was reached with individuals of zero complexity. Furthermore, we have not detected any similar

growth of complexity for any other ϵ value. To sum up, what has been observed is an intermediate behaviour between the two cases previously studied. There is induced complexity, although smaller than the $\mu = 2$ case. Smaller because of the difference between the individual complexity and the induced complexity, and smaller because complexity is not high except at the boundary between turbulent and ordered phases.

4.5 Conclusion

In this chapter we have analysed some computational properties of GCM. Our interest was to explore the existence of collective-induced computation in some natural systems (as ant colonies) where the single units behave very simply in isolation and in a complex way when forming part of the entire system. More precisely, we should ask how ant colonies formed by rather simple individuals (when isolated) can be able to induce them to perform complex computation, as observed.

The information-theoretic characterization of the phase space has shown that the Markov partition defined on the logistic map provides an adequate characterization. Information transfer, in particular, shows three different types of behavior: it is high at the coherent and ordered phases, close to zero at the turbulent regime and it takes intermediate values for glassy dynamics.

These quantities change rather sharply at the boundaries between different phases. This makes some difference in relation with previous studies, where information transfer becomes maximum at the phase transition (where correlations diverge). GCM do not show this type of maximum because of the globally coupled nature of the interactions. But for the same reason we expect to find some generic, common properties (in terms both of computation and dynamical properties) at each phase.

The ϵ-machine reconstruction of single maps close to the onset of chaos gives us a finite automaton with many states (here 31). So at this point we have a complex object in terms of computation. Interestingly, the coupling with other units via GCM do not modify this complexity. So entities which are computationally complex in isolation do not change in the presence of coupling: nothing new is induced by the collective. This observation matches the behavior of weakly evolved, primitive ants, where individuals are enough complex to work in isolation and the interactions among them are rather irrelevant.

However, if we start with random, computationally trivial maps and then couple them, the situation turns out to be very different. At $\mu = 2.0$ a fully chaotic map is obtained. The partition of this chaotic attractor defines a Bernouilli sequence and so

we have a $C_\mu = 0$ complexity. Starting from low couplings, at the turbulent domain, the reconstruction algorithm does not converge, as expected given the disordered, high-dimensional nature of the attractors. In spite of the remaining coherence no finite machines are obtained.

But as we reach the boundary between the turbulent and the ordered phases, the situation changes radically. Now the coherent motion and the spontaneous emergence of clustering also gives birth to well defined ϵ-machines. Suddenly, the coupling starts to control the dynamics of individuals and they behave in a computationally complex way. Nothing except the coupling has been introduced, but it is enough to generate complexity. As in the real ant colonies discussed in section 4.4.1, simple isolated individuals can behave in a complex way inside the collective. This is precisely what we have observed. A very important suggestion emerging from this result is that in collectives of simple agents complex behavior is only defined at the level of individuals *inside* the colony and not as isolated entities. In this sense, the observed behavior is the result of an emergent property.

Results in this chapter have been published in Delgado & Solé (1997b) and Delgado *et al.* (1997)

Chapter 5

Overview and prospects

5.1 Summary

It is a working hypothesis that either artificial or natural collective systems process (store, manipulate and transmit) information. It is not difficult to agree with this hypothesis; a close look at any ant nest make us feel sure of its truthfulness. Much work has been done to ascertain which are the hidden mechanisms underlying this fascinating behavior, work that has been undoubtedly fruitful, as we can see in the comprehensive treatise "The Ants" by Hölldobler & Wilson (1990). However, much work remains to be done. On one hand, we currently see a landscape of ad-hoc algorithms and robots, built in their majority with a stigmergic inspiration; on the other hand there is a plethora of theoretical models of computation, with some features that make us call them "collective" systems, that are too theoretically oriented to be useful when trying to understand real -physical- collective systems. The middle point, that of exploring real systems with computational notions, is almost a no-man's land. In fact, the two objections mentioned above have been the subject of this thesis: A first part trying to answer to "is there any alternative to stigmergic mechanisms when designing artificial systems?" and a second part, not totally disconnected from the first (see chapter 3) looking for the computational side of physical collective systems. These open problems, belonging to theoretical myrmecology and to RDAI, have been explored in this thesis. Self-synchronization has been analyzed in order to get *a)* a better understanding of the phenomenon and *b)* some other design principle of artificial collective systems. Then, the much less known field of computation with real physical collective systems has been explored with new methods: A NOR gate has been built with FNNs, though this approach was not completely adequate in order to attain our goal of relating collective dynamics with computation, nevertheless the result is quite robust, property not found in

previous dynamical systems behaving as such gates. Once left the FNN framework, the computational mechanics paradigm has been applied to the formalisation of an interesting problem in collective computation. Let us review our work, detailing the new results introduced in this thesis.

Fluid Neural Networks

The Fluid Neural Network (FNN) was originally created to model synchronization of activity in ant colonies. Previous work on density dependence of oscillations has shown that there is a critical density signaling the appareance of ordered temporal patterns of activity, and this critical density is close to that observed in real ant colonies. This fact would suggest that ant colonies live in a sort of critical state. FNNs are a quite general model and might be analysed emphasizing quite different aspects of their dynamics. We started our study focusing on the relation between the noise perturbing the system and the self-organized patterns of behaviour. First of all we have performed an analysis of individual stochastic behavior deriving the probability of having n spontaneous activations in a time interval T. This quantity clearly shows the absence of periodicity in individual behavior making evident the collective nature of the phenomena under study. The numerical analysis performed over the entire system shows that noise induced transitions in FNNs underlie the appareance of oscillations. To proceed to a more in-depth analysis, some simplifications of the original model have been done and a mean field theory of FNNs has been proposed. Defining the Simple FNN (SFNN), a simple approximation of the dynamics has been performed giving an explanation of the oscillatory behaviour. Next we have changed to more practical concerns exploring how a sort of abstract task may be distributed in a collective of simple agents. Our result is that the system performance is comparable to that of a system with permanently active individuals. Summarizing, our results on FNNs are

- Analysis of stochastic individual behaviour
- Characterization of the disorder-order transition as a noise induced transition
- Mean field theory of FNNs:
  - A new model, the Simple FNN
  - Analytical determination of SFNNs critical densities
- Numerical analysis of task distribution in a system with self-synchronized activity, and comparisons with systems displaying different temporal patterns of activity.

Collective Computation

The possibility of a noise induced mechanism behind the computational behviour of collective systems (ant colonies display functional behavior related with self-synchronization) made us to switch our attention to the relation between collective systems and computation. Since dynamics and computation (information processing) are intimately related in real physical collective systems, this information processing must be measured and/or modeled in some not yet well established way. Furthermore, provided there are good models of the dynamics of collective systems, our goal is to extract computational information about the real system from its dynamical behaviour through its dynamical model. Of the diverse possibilities of relating dynamical systems and computation theory, our choice has focused on two alternatives.

First we have explored the possibility of computing within the FNNs framework. We have built a NOR gate with a modified FNN with which, in principle, any Turing machine may be constructed, since the NOR gate is an universal gate. But this approach had some drawbacks, such as the difficulty in connecting gates and that logical circuits are a non-uniform model of computation. We have, then, left the FNN framework and turned to different approaches to computation and dynamical systems: the computational mechanics approach. Within this framework, it is a working hypothesis that it is impossible to a dynamical system not to compute, since any dynamical system processes information. Our focus on collective systems has made us to emphasize an open problem often stated in the entomologists community but never formalised before: is there a complex threshold beyond which no emergent behaviour is possible? We also believe that this question may be a question that concerns all collective systems in general. An adequate collective of dynamical systems with which to study this question is the Globally Coupled Map (GCM) since it is a mean field version of a huge class of systems, called Coupled Map Lattices, with which several interesting natural phenomena have been modeled, such as immune or neural networks. Besides, the individuals composing the GCM are logistic maps, discrete-time continuous-state systems representative of any system with a period-doubling route to chaos, so then, a quite general class of dynamical systems. Our result, analyzing from a computational mechanics point of view the GCMs, is that the collective is unable of inducing any behaviour in complex individuals and that this is not the case with simple individuals, that are able to show more complex behaviour within the collective. Summarizing:

- A new FNN with "chemical" communication has been proposed.

- A NOR logical gate has been built, with the new property of robustness against perturbations, property that other systems behaving similarly do not possess.

and

- A new characterization of the GCMs phase space in terms of information theoretic measures.
- A new way to deal with dynamics and computation in collective systems, applying the computational mechanics hypothesis to dynamical systems with a large number of individual components.
- A first approach to the formalisation of the problem concerning the relation "individual complexity/collective complexity" and a theoretical evidence of the existence of an upper bound in individual complexity beyond which the collective behaviour is no more than the sum of the individual behaviors.

5.2 Prospects

Collective systems with emergent functional properties, those of interest to RDAI and to a computational view of collective dynamical systems, are just beginning to reveal their secrets to physicists, biologists and computer scientists. In fact, the creation of the new and promising science called "Complexity" had to do with the growing interest in collective systems. These systems require notions from quite diverse fields, such as myrmecology, dynamical systems theory, statistical physics or computation theory to be fully understood, so they are a tangible example of the urgent need of interdisciplinariety in science, need that was claimed long time ago by N. Wiener in his classic "Cybernetics". The available results on collective systems have been reviewed in this thesis, showing that by no means there exists a "theory of emergent collective systems", but a large collection of results that are not, at the moment, general enough. There are still a lot of open questions on collective systems.

The work on FNN may be seen as a step forward in the analysis of collectives of *mobile* automata, though it is not clear if these systems form a unified class of systems. The work of Boccara *et al.* (1994) classifies a mobile cellular automata in the universality class of directed percolation (Binney *et al.*, 1992) though the same cannot be stated about FNNs, which display a *finite-size* transition, that is, in the limit $L \to \infty$ there is no self-organized oscillations. This may be clearly understood from the mean-field analysis of SFNN, since the time of activity propagation would

be infinite while the time of individual inactivation would be finite, preventing any synchronization. However this idea must be sistematically developed and related with other work on finite-size transitions:

- Finite-size transitions of FNNs

Another poorly explored aspect of FNNs is the effect of the connections J_{ij} on the dynamics. Fortunately we have measures on real *Leptothorax* colonies and a lot of work may be done with the empirically estimated J_{ij} (Cole & Cheshire, 1996).

- Dynamics of FNNs with "real" J_{ij}

The critical densities in simple FNN (SFNN) have been analytically computed, though the approximation of τ_{per} is not completely satisfactory. A more sound analytical determination of τ_{per} would be quite desirable:

- Better analytical calculation of τ_{per}

Perhaps the less explored side of FNNs is that of their practical consequences. We have shown that self-synchronization saves work to individuals allowing a great efficiency, though some more real application would be of interest to definitely consider self-synchronization as an alternative to stigmergy. There are some immediate possibilities, such as measure how well allocated are tasks in self-synchronized systems, following some natural clues to design the task distribution, in particular, a concentric task distribution in the space of the colony (Franks *et al.*, 1992 and Guy Theraulaz, personal communication):

- Explore the performance of FNNs with schemes of task distribution based on real data.

These are some immediate questions to proceed with the work on FNNs, but there are interesting questions on collective systems that may be answered within the framework of FNNs. One of these, the existence of phase transitions in social insect behavior, has been dealt with before (Oster & Wilson, 1978; Bonabeau, 1996; Millonas, 1992, 1994) but always in models not taking into account synchronization of temporal activity. We are quite sure that one of the possibilities to include self-synchronization in this work is in the utilization of FNN-like systems.

Computation in collective systems is a relatively new field and definitive results must yet to appear, since current theories are not satisfactory, as has been discussed at length in the introduction and in chapters 3 and 4. Our goal is to study physical collective systems with emergent functionality and our compromise is with the

"computational mechanics" approach. With respect to a general theory, much more models of collective systems must be studied with the same techniques we have applied to globally coupled maps to assure the suitability of our approach. The same applies to future work on the problem we have dealt with, that of "individual complexity/collective complexity":

- Generalization of our computational mechanics approach to GCMs to other models of collective systems.

However, the generalization must be done with some care, since the globally coupled maps are representative of a huge class of collective dynamical systems (Kaneko, 1990a, 1990b). We may also proceed with the study of the "individual complexity/collective complexity" within the globally coupled maps framework. The idea is the following: could we measure how the interaction with the collective is behaving with respect to the individual? We may formalize this idea measuring the automaton associated to one individual of the GCM (chosen at random) and at the same time computing an histogram of the mean state of the entire system. Later, this histogram may be used as the probability density of a noise perturbing a logistic map, and computing the automaton associated to this noisy map (James P. Crutchfield, personal communication). Let us recall the GCM

$$x_{n+1}(i) = (1-\epsilon) f_\mu(x_n(i)) + \frac{\epsilon}{N} \sum_{j=1}^{N} f_\mu(x_n(j)) \tag{5.1}$$

Then, from the temporal evolution of the system an histogram of

$$\frac{1}{N} \sum_{j=1}^{N} f_\mu(x_n(j)) \tag{5.2}$$

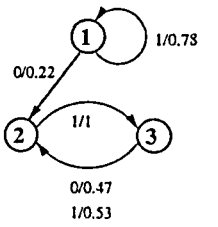
is computed and it will be used to compute the probability distribution of ξ_n, a noise perturbing the system

$$x_{n+1} = (1-\epsilon) f_\mu(x_n) + \epsilon \xi_n \tag{5.3}$$

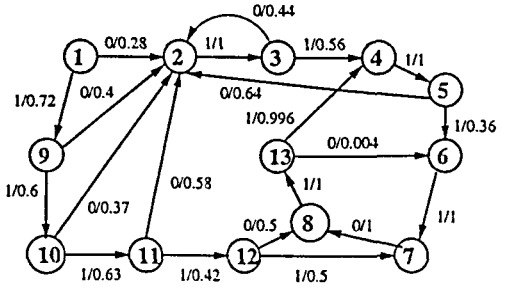
Furthermore, we are working now with μ parameters such that the metric entropy h_μ of the logistic map

$$x_{n+1} = f_\mu(x_n)$$

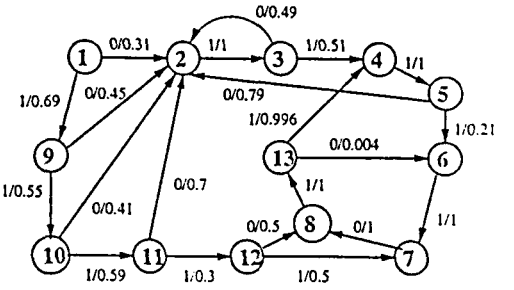
is $0 < h_\mu < 1$, that is, neither complete order nor complete disorder. In the results introduced in chapter four, all the automata had either $h_\mu = 0$ or $h_\mu = 1$, in order to simplify the work. Our preliminary results are quite surprising, since in the turbulent



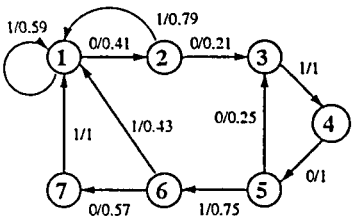
Logistic map $\mu = 1.5436$
h = 0.53
C = 1



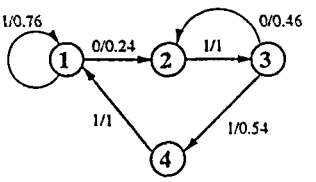
Globally Coupled Map
h = 0.40 $N = 500$
C = 2.81 $\mu = 1.5436$
$\varepsilon = 0.025$



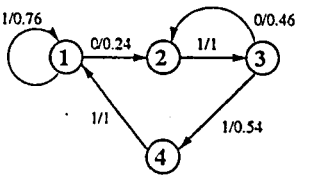
Logistic map with noise
h = 0.41 $N = 500$
C = 2.56 $\mu = 1.5436$
$\varepsilon = 0.025$



Logistic map $\mu = 1.85$
h = 0.76
C = 2.06

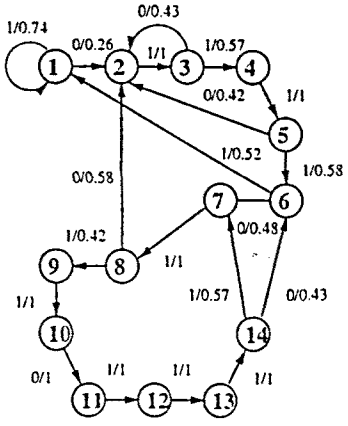


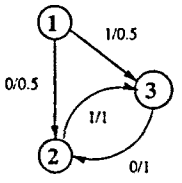
Globally Coupled Map
h = 0.59 $N = 500$
C = 1.82 $\mu = 1.85$
$\varepsilon = 0.1$

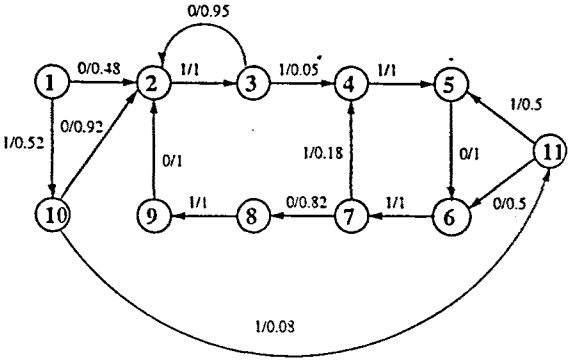


Logistic map with noise
h = 0.59 $N = 500$
C = 1.82 $\mu = 1.85$
$\varepsilon = 0.1$

Figure 5.1: The three automata obtained from the logistic map, a randomly chosen individual of the GCM and a map with a noisy perturbation, with parameters μ and ϵ such that the GCM belongs to the turbulent region of the phase space (see figure 4.7) (C is C_μ and h is h_μ)



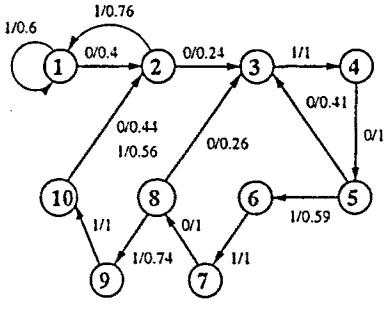


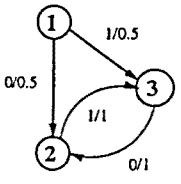


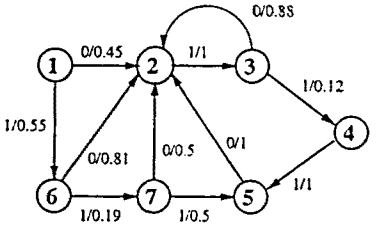
Logistic map μ=1.6
h = 0.58
C = 3.29

Globally Coupled Map
h = 0 *N = 500*
C = 1 μ = *1.6*
ε = *0.15*

Logistic map with noise
h = 0.17 *N = 500*
C = 1.77 μ = *1.6*
ε = *0.15*







Logistic map μ=1.8
h = 0.66
C = 2.95

Globally Coupled Map
h = 0 *N = 500*
C = 1 μ = *1.8*
ε = *0.2*

Logistic map with noise
h = 0.26 *N = 500*
C = 1.49 μ = *1.8*
ε = *0.2*

Figure 5.2: The three automata obtained from the logistic map, a randomly chosen individual of the GCM and a map with a noisy perturbation, with parameters μ and ϵ such that the GCM belongs to the ordered region of the phase space (see figure 4.7). In this case, the GCM automaton is different from the noisy automaton. (C is C_μ and h is h_μ)

region of the phase space (see figure 4.7) the automata of the individual chosen at random from the GCM and that of the noisy map are the same automaton, though this is not the case in the ordered region (see figures 5.1 and 5.2) so it looks like as if we could characterize the phase space from the obtained automata. What does this mean? In principle, the first conclusion we can suggest is that, *computationally*, the interaction of the collective with the individual is indistinguishable from a noise in the turbulent region, whereas in the ordered case the interaction possesses some structure, as reflected in the more ordered behaviour of the GCM individual when compared with the noisy individual. Nevertheless, this work is still in a preliminary stage and more data and/or analysis is required to draw any definitive conclusions. So then

- Characterization of the GCM phase space with statistical complexity measures and quantification of the interaction of the individual with the collective by means of finite automata.

Of course, some research to clarify the Crutchfield's conjecture (see chapter 4) will be of great utility to go on with the computational view of collective systems:

- Is $C_{us}^{A}(I) \leq C_{in}^{A}(I)$ in general?

Finally, our results on a theory of collective computation are the first steps towards a systematic new view of collective systems in terms of computing structures. This is part of a large project to see nature with computational "glasses", with the certainty that this approach will provide us with useful information on physical processes. Of course "standard" computation theory does not suffice, it must be enlarged to deal with space (in a dynamical systems sense), probability and continuity and the steps in this direction are currently an active research topic. Our choice of DFAs as the computational structure to work with deserves a final comment. Until now only the process of reconstructing finite automata is fully automated, which seems to be, at first sight, insufficient considering that DFA are at the bottom of the Chomsky Hierarchy. But some recent results on noisy computation could make us change that view. There are some models of analog computation that are universal, in the sense that they can simulate an universal Turing machine, but if we introduce some quantity of noise the computational power of those models is reduced to that of DFA (Maas & Orponen, 1996 and references therein). On the other hand, noise is ubiquitous in nature, as we have been discussing in this thesis, therefore all good computational models of biological phenomena have to take into account that noise, so, could it be that the computational power of DFA is enough when modeling nature? Of course the question is completely open.

In the meantime, let us quote J.P. Crutchfield, answering to the question *how can we detect and then quantify structure in natural processes?*:

> *In pursuing answers to this sort of question we've come to the conclusion that the diverse model classes found in computation theory are key tools in being explicit about how natural information processing mechanisms can be represented and analyzed. However, we also have come to the conclusion that contemporary notions of "computation" and of "useful" information processing -colored as they are by the recent history of digital computer technology- must be extended in order to be useful within empirical science. Why? Because the processes studied by natural scientists involve systems that are continuous, stochastic, spatially extended, or some combination of these and other characteristics that fall strictly outside the purview of discrete computation theory.*

Chapter 6

References

ADLER, F.R. & GORDON, D.M. (1992) Information collection and spread by networks of patrolling ants. *Am. Nat.* **140**, 373-400.

ALLOWAY, T.M., BUSCHINGER, A., TALBOT, M., STUART, R., & THOMAS, C. (1982) Polygyny and polydomy in three north american species of the ant genus Leptothorax (Mayr). *Psyche* **89**, 249-274.

AMIT, D. (1989). *Modeling Brain Function.* Cambridge: Cambridge University Press.

ASH, R.B. (1965). *Information Theory.* New York: Dover

BAETEN, J.C.M. & BERGSTRA (1991). Real space process algebra. *Formal Aspects of Computing* **3**, 142-188

BALCÁZAR, J.L., DÍAZ, J. & GABARRÓ, J. (1988-1990). *Structural Complexity, vols. I and II.* EATCS Monographs. Berlin: Springer-Verlag

BECK, C. & SCHLÖGL, F. (1993) *Thermodynamics of Chaotic Systems.* Cambridge: Cambridge University Press

BECKERS, R., DENEUBOURG, J.L., GOSS, S., & PASTEELS, J.M. (1990) Collective decision making through food recruitment. *Ins. Soc.* **373**, 258-267.

BEER, R.D. (1990) *Intelligence as Adaptive Behavior*, San Diego: Academic Press.

BEER, R.D. (1995) A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence* **72**, 173-215.

BENI, G. & HACKWOOD, S. (1992), The maximum entropy principle and sensing in swarm intelligence. In *Towards a practice of autonomous systems* (Varela, F. & Bourgine, P. eds.), pp. 153-160, Cambridge, Massachusetts: MIT Press.

BERLEKAMP, E.R., CONWAY, J.H. & GUY, R.K. (1985) *Winning ways for your mathematical plays, vol. 2*, London: Academic Press.

BINNEY, J.J., DOWRICK, N.J., FISHER, A.J. & NEWMAN, M.E.J. (1992) *The theory of critical phenomena: An introduction to the renormalization group*, Oxford: Oxford Univ. Press.

BOCCARA, N., KYEONG, C. & ORAM, M. (1994). A probabilistic automata network epidemic model with births and deaths exhibiting cyclic behaviour. *J. Phys. A: Math. Gen.* **27**, 1585-1597.

BONABEAU, E. & THERAULAZ, G. (eds.) (1995). *Intelligence Collective.* Paris: Hermes

BONABEAU, E (1996) Marginally stable swarms are flexible and efficient. *J. Phys. I France* , **6**, 309-324

BONABEAU, E (1997) Flexibility at the edge of chaos: a clear example from foraging in ants. *Acta Biotheoretica* (in press)

BONABEAU, E., THERAULAZ, G. & DENEUBOURG, J.L. (1996). Quantitative study of the fixed threshold model for the regulation of division of labour in insect societies. *Proc. R. Soc. Lond. B* **263**, 1565-1569.

BONABEAU, E., THERAULAZ, G. &, DENEUBOURG J.L. (1997) Synchronization and foraging efficiency in ants. *Preprint*

BOND, A.H. & GASSER, L. (1988), *Readings in Distributed Artificial Intelligence*, California: Morgan Kaufmann Publishers.

BRANICKY, M.S. (1995). Universal computation and other capabilities of hybrid and continuous dynamical systems. *Theor. Comp. Sci.* **138**, 67-100.

BROOKS, R.A. (1991). Intelligence without representation, *Artificial Intelligence* **47**, 139-159.

CASTI, J.L. (1992) *Reality Rules: Picturing the World with Mathematics. Vol. 2, the frontier.* New York: John Wiley & Sons.

CASTILLO, L. & QUINTANILLA, G. (1991). Initial concepts in Distributed Artificial Intelligence. In *Proceedings de TEC-COMP 91, Computación No Convencional: Hacia los Sistemas Inteligentes*, Apr. 1991, Mexico D.F.

CECCATTO, H.A. & HUBERMAN, B.A. (1989). Persistence of nonoptimal strategies. *Proc. Natl. Acad. Sci. USA* **86**, 3443-3446.

CHATÉ, H. & MANNEVILLE, P. (1989). Coupled map lattice as cellular automata. *J. of Stat. Phys.* **56**, 357-370.

COLE, B.J. (1986). The social behavior of *Leptothorax allardycei* (*Hymenoptera, Formicidae*) - time budgets and the evolution of worker reproduction. *Behav. Ecol. Sociobiol.* **18**, 165-173.

COLE, B.J. (1991a). Short-term activity cycles in ants: generation of periodicity by worker interaction. *Am. Nat.* **137**, 244-259.

COLE, B.J. (1991b). Is animal behaviour chaotic? Evidence from the activity of ants. *Proc. R. Soc. Lond. B* **244**, 253-259.

COLE, B.J. & CHESHIRE, D. (1996). Mobile cellular automata models of ant behavior: movement activity of *Leptothorax Allardycei. Am. Nat.* **148**, 1-15.

COLORNI, A., DORIGO, M. & MANIEZZO, V. (1992). An investigation of some properties of an Ant algorithm. In *Parallel Problem Solving from Nature 2* (Männer, R. & Manderick, B. eds.) pp. 509-520. Elsevier Science.

CRUTCHFIELD, J.P. (1994a) Critical computation, phase transitions and hierarchical learning. In *Towards the Harnessing of Chaos* (Yamaguti, M. ed.). Amsterdam: Elsevier Science.

CRUTCHFIELD, J.P. (1994b) The calculi of emergence: Computation, dynamics and induction. *Physica D* **75**, 11-54

CRUTCHFIELD, J.P. & PACKARD, N.H. (1983) Symbolic dynamics of noisy chaos. *Physica D* **7**, 201-223.

CRUTCHFIELD, J.P. & YOUNG, K. (1989) Inferring statistical complexity. *Phys. Rev. Lett.* **63**, 105-108

CRUTCHFIELD, J.P. & YOUNG K. (1990) Computation at the onset of chaos. In *Entropy, Complexity and Physics of Information* (Zurek, W. ed.) pp. 223-269. Reading, Massachusetts: Addison-Wesley.

CRUTCHFIELD, J.P. & FELDMAN, D.P. (1997) Statistical complexity of simple one-dimensional spin systems. *Phys. Rev. E* **55**, 1239R-1244R

DAVIS, R. & SMITH, R.G. (1983) Negotiation as a metaphor for distributed problem solving. *Artificial Intelligence* **20**, 63-109

DELGADO, J. & SOLÉ, R.V. (1997a) Noise Induced Transitions in Fluid Neural Networks. *Phys. Lett. A* **229**, 183-189.

DELGADO, J. & SOLÉ, R.V. (1997b) Collective-induced computation. *Phys. Rev. E* **55**, 2338-2344.

DELGADO, J. & SOLÉ, R.V. (1997c) Mean field theory of Fluid Neural Networks. Research Report LSI-97-27-R, *Submitted to Phys. Rev. E*

DELGADO, J. & SOLÉ, R.V. (1997d) Self-synchronization and task fulfilment in ant colonies. Santa Fe Institute Working Paper, *Submitted to Proc. Roy. Soc. B*

DELGADO, J., SOLÉ, R.V., GOODWIN, B.C. & CRUTCHFIELD, J.P. (1997) Computation and emergence in ant colonies. *Submitted to J. Theor. Biol.*

DENEUBOURG, J.L. (1977) Application de l'ordre par fluctuations a la description de certaines étapes de la construction du nid chez les termites. *Ins. Soc.* **24**, 117-130.

DENEUBOURG, J.L., GOSS, S., FRANKS, N. & PASTEELS, J.M. (1982) The blind leading the blind: Modeling chemically mediated army ant raid patterns. *J. Insect Behav.* **2**, 719-725.

DENEUBOURG, J.L., PASTEELS, J.M. & VERHAEGHE, J.C. (1983) Probabilistic behavior in Ants: A strategy of errors? *J. Theor. Biol.* **105**, 259-271.

DENEUBOURG, J.L., ARON, S., GOSS, S., PASTEELS, J.M. & DUERINCK, G. (1986) Random behavior, amplification processes and number of participants: How they contribute to the foraging properties of ants. *Physica D* **22**, 176-186.

DENEUBOURG, J.L., ARON, S., GOSS, S. & PASTEELS, J.M. (1990) Self-organizing exploratory pattern of the Argentine Ant. *J. Insect Behav.* **32**, 159-168.

DENEUBOURG, J.L., GOSS, S., FRANKS, N., SENDOVA-FRANKS, A., DETRAIN, C. & CHRÉTIEN, L. (1991) The dynamics of collective sorting, robot-like ants and ant-like robots. In *From Animals to Animats* (Meyer, J.A. & Wilson, S.W. eds.) pp. 356-363. Cambridge, Massachusetts: MIT Press.

DENEUBOURG, J.L., THERAULAZ, G. & BECKERS, R. (1992), Swarm-made architectures. In *Towards a practice of autonomous systems* (Varela, F. & Bourgine, P. eds.), pp. 123-133, Cambridge, Massachusetts: MIT Press.

DORIGO, M., MANIEZZO, V., & COLORNI, A. (1996) The Ant System: Optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man and Cybernetics-Part B* **26**, 29-41.

DREYFUS, H.L. (1992), *What computers still can't do: A critique of artificial rea-*

son, Cambridge: MIT Press.

FORREST, S. (1990) Emergent computation: self-organizing, collective and cooperative phenomena in natural and artificial computing networks. *Physica D* **42**, 1-11.

FRANKS, N.R. & BRYANT, S. (1987) Rhythmical patterns of activity within the nest of ants. In *Chemistry and Biology of Social Insects* (Eder, J. & Rembold, H. eds.) pp. 122-123. Munich: Peperny.

FRANKS, N.R., BRYANT, S., GRIFFITHS, R. & HEMERIK, L. (1990) Synchronization of the behavior within the nests of the ant *Leptothorax acervorum*-I. Discovering the phenomenon and its relation to the level of starvation. *Bull. Math. Biol.* **52**, 597-612.

FRANKS, N.R., WILBY, A., SILVERMAN, B.W. & TOFTS, C. (1992) Self-organizing nest construction in ants: Sophisticated building by blind bulldozing. *Anim. Behav.* **44**, 357-375.

GACS, P. & REIF, J. (1988) A simple three-dimensional real-time reliable cellular array *J. Comput. System Sci.* **36**, 125-147.

GARDINER, C.W. (1983) *Handbook of stochastic methods* Berlin: Springer-Verlag

GARZON, M. (1995) *Models of Massive Parallelism*, Berlin: Springer-Verlag.

GOODWIN, B. (1995) *How the Leopard Changed its Spots* London: Weidenfeld and Nicolson.

GORDON, D.M. (1989a) Caste and change in social insects. In *Oxford Surveys in Evolutionary Biology* (Harvey, P.H. & Partridge, L. eds), Vol.6

GORDON, D.M. (1989b) Dynamics of task switching in harvester ants *Anim. Behav.* **38**, 194-204.

GORDON, D.M., GOODWIN, B.C. & TRAINOR L.E.H. (1992) A parallel distributed model of the behaviour of ant colonies *J. Theor. Biol.* **156**, 293-307

GOSS, S. & DENEUBOURG, J.L. (1992), Harvesting by a group of robots. In *Towards a practice of autonomous systems* (Varela, F. & Bourgine, P. eds.), pp. 195-204, Cambridge, Massachusetts: MIT Press.

GUTOWITZ, H.A. (1993) Complexity-seeking ants. *Proceedings of the third European conference on Artificial Life*

HAKEN, H. (1977) *Synergetics* Berlin: Springer-Verlag

HAKEN, H. (1988) *Information and Self-organization* Berlin: Springer-Verlag

HANSON, J.E. (1993) *Computational Mechanics of Cellular Automata*, Ph. D. Dissertation, University of California, Berkeley.

HATCHER, M.J., TOFTS, C. & FRANKS, N.R. (1992) Mutual exclusion as a mechanism for information exchange within ant nests *Naturwissenschaften* **79**, 32-34.

HERBERS, J.M. (1983) Social organization in Leptothorax ants: Within- and between-species patterns *Psyche* **90**, 361-386.

HOFSTADTER, D. (1979), *Gödel, Escher, Bach*, Brighton: Harvester Press.

HOGG, T. & HUBERMAN, B.A. (1987) Artificial intelligence and large scale computation: A physics perspective. *Phys. Rep.* **156** 227-310.

HÖLLDOBLER, B. & WILSON, E.O. (1990) *The Ants* Berlin: Springer-Verlag

HOPCROFT, J.E. & ULLMAN, J.D. (1979) *Introduction to Automata Theory, Languages and Computation* Reading, Massachusetts: Addison-Wesley.

HORSTHEMKE, W. & LEFEVER, R. (1984) *Noise-Induced Transitions*, Berlin: Springer-Verlag.

HUBERMAN, B.A. (ed.) (1988) *The Ecology of Computation*, North-Holland

KANEKO, K. (1990a) Clustering, coding, switching, hierarchical ordering and control in a network of chaotic elements *Physica D* **41**, 137-172.

KANEKO, K. (1990b) Globally coupled chaos violates the law of large numbers but not the central-limit theorem *Phys. Rev. Lett.* **65**, 1391-1394

KANEKO, K. (1994). Relevance of dynamic clustering to biological networks. *Physica D* **75**, 55-73

KELSO, J.A.S. (1995). *Dynamic Patterns: The Self-Organization of Brain and Behavior.* Cambridge: MIT Press.

KEPHART, J.O., HOGG, T. & HUBERMAN, B.A. (1990), Collective Behavior of predictive agents. *Physica D* **42**, 48-65.

KUBE, R. & ZHANG, H. (1993), Collective robotics: From social insects to robots, *Adaptive Behav.* **2**, 189.

LÉVI-STRAUSS, C. (1962), *La Pensée Sauvage*, Paris: Plon

LUENBERGER, D.G. (1979), *Introduction to Dynamic Systems*, John Wiley & Sons.

MAAS, W. & ORPONEN, P. (1996) On the effect of analog noise in discrete-time analog computations. NeuroCOLT Technical Report NC-TR-97-021.

MAETERLINCK, M. (1927), *The Life of the White Ant*, London: Allen and Unwin.

MATARIC, M.J. (1994), Designing emergent behaviors: From local interactions to collective intelligence. In *Artificial Life IV* (Brooks, R.A. & Maes, P. eds.), pp. 432-441, Cambridge, Massachusetts: MIT Press.

McCORDUCK, P. (1979), *Machines who think*, W.H. Freeman & Co.

MIKHAILOV, A.S. (1990), *Foundations of Synergetics I*, Berlin: Springer-Verlag.

MIKHAILOV, A.S. (1993), Collective dynamics in models of communicating populations. In *Interdisciplinary Approaches to Nonlinear Complex Systems. Springer Series in Synergetics vol.62* (Haken, H. & Mikhailov, A. eds.) pp.89-108. Berlin: Springer Verlag.

MILLONAS, M.M. (1992) A connectionist type model of self-organized foraging and emergent behavior in ant swarms *J. theor. biol.* **159**, 529-552.

MILLONAS, M.M. (1994) Swarms, Phase Transitions and Collective Intelligence. In: *Artificial Life III* (Langton, C. ed.) pp 417-445. Addison-Wesley

MILNER, R. (1990) *Communication and Concurrency*, Englewood, Cliffs: Prentice-Hall

MIRAMONTES, O. (1992) *Complexity and Behaviour in Leptothorax Ants*, Master Thesis, Open University (U.K.)

MIRAMONTES, O. (1995) Order-disorder transitions in the behavior of ant societies. *Complexity* **1**(2), 56-60.

MIRAMONTES, O., SOLÉ, R.V. & GOODWIN, B.C. (1993) Collective behaviour of random-activated mobile cellular automata. *Physica D* **63**, 145-160.

MITCHELL, M., CRUTCHFIELD, J.P. & HRABER P.T. (1994) Evolving cellular automata to perform computations: Mechanisms and impediments. *Physica D* **75**, 361-391.

MORITZ, R.F.A. & SOUTHWICK, E.E. (1992) *Bees as a superorganism, an evo-*

lutionary reality Berlin: Springer-Verlag.

MURPHY, K.P. (1996) Passively Learning Finite Automata, SFI Working Paper SFI-96-04-017, Santa Fe Institute, Santa Fe, New Mexico, USA

NICOLIS, G. & PRIGOGINE, I. (1977) *Self-organization in non-equilibrium systems* New York: Wiley

NICOLIS, G. & PRIGOGINE, I. (1989) *Exploring Complexity* New York: Freeman.

OSTER, G.F. & WILSON, E.O. (1978) *Caste and ecology in the Social Insects* Princeton: Princeton University Press.

PACCALA, S.W., GORDON, D.M. & GODFRAY, H.C.J. (1996) Effects of social group size on information transfer and task allocation *Evolutionary Ecology* **10**, 127-165.

PEITGEN, H.-O., JÜRGENS, H. & SAUPE, D. (1992). *Chaos and Fractals: New Frontiers of Science.* New York: Springer-Verlag.

RICCIARDI, L.M. (1977) *Diffusion processes and related topics in biology.* Lecture notes in biomathematics vol. 14, Berlin: Springer Verlag

ROBINSON, G.E. (1992) Regulation of division of labor in insect societies *Ann. Rev. Entomol.* **37**, 637-702.

RUSSELL, S. & NORVIG, P. (1995) *Artificial intelligence: A modern approach*, New Jersey: Prentice-Hall

SCHUSTER, H.G. (1989). *Deterministic Chaos: An Introduction*, 2nd ed. Weinheim: VCH.

SENDOVA-FRANKS, A. & FRANKS, N.R. (1993) Task allocation in ant colonies within variable environments (a study of temporal polyethism: experimental). *Bull. Math. Biol.* **55**, 75-96.

SMITHERS, T. (1994) What the dynamics of adaptive behaviour and cognition might look like in agent-environment interaction systems. In *Proceedings of DRABC'94, III International Workshop on Artificial Life and Artificial Intelligence*, Euskal Herriko Unibersitatea, pp. 134-153.

SOLÉ, R.V. & DELGADO, J. (1996) Universal Computation in Fluid Neural Networks *Complexity* 2(2), 49-56

SOLÉ, R.V., MANRUBIA, S.C., LUQUE, B., DELGADO, J. & BASCOMPTE, J.

(1996) Phase Transitions and Complex Systems *Complexity* **1**(4), 13-26.

SOLÉ, R.V. & MIRAMONTES, O. (1995) Information at the edge of chaos in Fluid Neural Networks *Physica D* **80**, 171-180

SOLÉ, R.V., MIRAMONTES, O. & GOODWIN, B.C. (1993a) Oscillations and chaos in ant societies. *J. Theor. Biol.* **161**, 343-357.

SOLÉ, R.V., MIRAMONTES, O. & GOODWIN, B.C. (1993b) Emergent behavior in insect societies: Global oscillations, chaos and computation. In *Interdisciplinary Approaches to Nonlinear Complex Systems. Springer Series in Synergetics vol.62* (Haken, H. & Mikhailov, A. eds.) pp.77-88. Berlin: Springer Verlag.

STEELS, L. (1991) Towards a theory of emergent functionality. In *From Animals to Animats* (Meyer, J.A. & Wilson, S.W. eds.) pp. 451-461. Cambridge, Massachusetts: MIT Press.

SUDD, J.H. (1967) *An Introduction to the Behaviour of Ants* London: Arnold.

SUGAWARA, K. & SANO, M. (1997) Cooperative acceleration of task performance: Foraging behavior of interacting multi-robots system. *Physica D* **100**, 343-354.

TOFTS, C., HATCHER, M. & FRANKS, N.R. (1992) The autosynchronization of the ant *Leptothorax acervorum* (Fabricius): Theory, testability and experiment *J. Theor. Biol.* **157**, 71-82

TOFTS, C. (1993) Algorithms for task allocation in ants (a study of temporal polyethism: Theory) *Bull. Math. Biol.* **55**, 891-918

TOFTS, C. (1994) Processes with probabilities, priority and time *Formal Aspects of Computing* **6**, 536-565.

TOFTS, C. (1996) Describing social insect behaviour using process algebra *preprint* (http://www.scs.leeds.ac.uk/chris/Papers/dsib.ps.Z)

von NEUMANN, J. (1956) Probabilistic logics and the synthesis of reliable organisms from unreliable components. In *Automata Studies (Ann. of Math. Studies vol. 34)* (Shannon, C.E. & McCarthy, J. eds.) pp. 43-98. Princeton: Princeton University Press

WIGGINS, S. (1990) *Introduction to Applied Nonlinear Dynamical Systems and Chaos* Berlin: Springer-Verlag.

WILSON E.O. (1971) *The Insect Societies*, Cambridge, Massachusetts: The Belknap Press of Harvard University Press

WOOLDRIDGE, M.J. & JENNINGS, N.R. (eds.) (1995) *Intelligent Agents.* Lecture Notes in Artificial Intelligence, vol. 890. Berlin: Springer-Verlag.